DOCUMENT RESUME

ED 056 045 24 TE 002 649

AUTHOR	Houston, Thomas Rappe, Jr.
TITLE	Comparable Common Factors in English Homophone Recognition. Report from the Project on Situational Variables and Efficiency of Concept Learning.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	TR-163
BUREAU NO	BR-5-0216
PUB DATE	Mar 71
CONTRACT	OEC-5-10-154
NOTE	178p.
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	*Concept Formation; *English; Factor Analysis; *Language Research; Projects; *Tests; *Word Recognition

ABSTRACT

A homophone is a word having the same pronunciation as another English word, but a different spelling. A list of 7,300 English homophones was compiled and used to construct two tests. Scores were obtained in these and on reference tests for J. P. Guilford's factors CMU, CSU, DMU, and DSU for 70 native speakers of midwestern American English from a university population. The homophone tests showed Hoyt reliabilities of .95 and .87 for these subjects. Following Harris's procedure for determining Comparable Common Factors, the 15 x 15 matrix of intercorrelations was subjected to three factoring procedures, each yielding oblique and orthogonal solutions. Results were in close agreement for all analyses, yielding three common factors. Two corresponded to CMU and to DMU. The CSU and DSU tests loaded on the third factor, which had its largest loadings on the homophone tests, and involved comparing verbal stimuli with formal elements of internally generated lists. These findings replicate Harris's failure to extract distinct CSU and DSU factors and suggest that homophone recognition tasks can provide short but reliable reference tests for the symbolic factor into which CSU and DSU coalesce. (Author)

OE-NCERD
TE
BR5-0216
PA24

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED056045

TE 002 649

1

U.S. Office of Education Center No. C-03 Contract OE 5-10-154

Technical Report No. 163

COMPARABLE COMMON FACTORS IN ENGLISH HOMOPHONE RECOGNITION

By Thomas Rappe Houston, Jr.

Report from the Project on Situational Variables
and Efficiency of Concept Learning

Herbert J. Klausmeier
V.A.C. Henmon Professor of Educational Psychology
Robert E. Davidson, Joel R. Levin, Thomas A. Romberg, B. Robert Tabachnick,
Alan Voelker, Larry Wilder, Peter Wolff
Technical Development Program Director: Mary R. Quilling
Research Associate: Dorothy A. Frayer

Robert E. Davidson, Associate Professor of Educational Psychology
and Chairman of the Examining Committee

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin
March, 1971

This Technical Report is a doctoral dissertation reporting research supported by the Wisconsin Research and Development Center for Cognitive Learning. Since it has been approved by a University Examining Committee, it has not been reviewed by the Center. It is published by the Center as a record of some of the Center's activities and as a service to the student. The bound original is in The University of Wisconsin Memorial Library.

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

2/3

NATIONAL EVALUATION COMMITTEE

Samuel Brownell
Professor of Urban Education
Graduate School
Yale University

Launor F. Carter
Senior Vice President on
Technology and Development
System Development Corporation

Francis S. Chase
Professor
Department of Education
University of Chicago

Henry Chauncey
President
Educational Testing Service

Martin Deutsch
Director, Institute for
Developmental Studies
New York Medical College

Jack Edling
Director, Teaching Research
Division
Oregon State System of Higher
Education

Elizabeth Koontz
Wage and Labor Standards
Administration, U.S.
Department of Labor,
Washington

Roderick McPhee
President
Punahou School, Honolulu

G. Wesley Sowards
Director, Elementary Education
Florida State University

Patrick Suppes
Professor
Department of Mathematics
Stanford University

***Benton J. Underwood**
Professor
Department of Psychology
Northwestern University

RESEARCH AND DEVELOPMENT CENTER POLICY REVIEW BOARD

Leonard Berkowitz
Chairman
Department of Psychology

Archie A. Buchmiller
Deputy State Superintendent
Department of Public Instruction

Robert E. Grinder
Chairman
Department of Educational
Psychology

Russell J. Hosler
Professor, Curriculum
and Instruction

Clauston Jenkins
Assistant Director
Coordinating Committee for
Higher Education

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational
Psychology

Stephen C. Kleene
Dean, College of
Letters and Science

Donald J. McCarty
Dean
School of Education

Ira Sharkansky
Associate Professor of Political
Science

B. Robert Tabachnick
Chairman, Department
of Curriculum and
Instruction

Henry C. Weinlick
Executive Secretary
Wisconsin Education Association

M. Crawford Young
Associate Dean
The Graduate School

EXECUTIVE COMMITTEE

Edgar F. Borgatta
Brittingham Professor of
Sociology

Anne E. Buchanan
Project Specialist
R & D Center

Robin S. Chapman
Research Associate
R & D Center

Robert E. Davidson
Assistant Professor,
Educational Psychology

Frank H. Farley
Associate Professor,
Educational Psychology

Russell J. Hosler
Professor of Curriculum and
Instruction and of Business

***Herbert J. Klausmeier**
Director, R & D Center
Professor of Educational
Psychology

Wayne Otto
Professor of Curriculum and
Instruction (Reading)

Robert G. Petzold
Associate Dean of the School
of Education
Professor of Curriculum and
Instruction and of Music

FACULTY OF PRINCIPAL INVESTIGATORS

Vernon L. Allen
Professor of Psychology

Ted Czajkowski
Assistant Professor of Curriculum
and Instruction

Robert E. Davidson
Assistant Professor of
Educational Psychology

Gary A. Davis
Associate Professor of
Educational Psychology

M. Vere DeVault
Professor of Curriculum and
Instruction (Mathematics)

Frank H. Farley
Associate Professor of Educational
Psychology

Lester S. Golub
Lecturer in Curriculum and
Instruction and in English

John G. Harvey
Associate Professor of
Mathematics and of Curriculum
and Instruction

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational
Psychology

Donald Lange
Assistant Professor of Curriculum
and Instruction

James Moser
Assistant Professor of Mathematics
Education; Visiting Scholar

Wayne Otto
Professor of Curriculum and
Instruction (Reading)

Milton O. Pella
Professor of Curriculum and
Instruction (Science)

Thomas A. Romberg
Associate Director, R & D Center
Professor of Mathematics and of
Curriculum and Instruction

B. Robert Tabachnick
Chairman, Department
of Curriculum and
Instruction

Richard L. Venezky
Assistant Professor of English
and of Computer Sciences

Alan Voelker
Assistant Professor of Curriculum
and Instruction

Larry Wilder
Assistant Professor of Curriculum
and Instruction

Peter Wolff
Assistant Professor of Educational
Psychology

MANAGEMENT COUNCIL

Herbert J. Klausmeier
Director, R & D Center
V.A.C. Henmon Professor of
Educational Psychology

Mary R. Quilling
Director
Technical Development Program

Thomas A. Romberg
Associate Director

James Walter
Director
Dissemination Program

Dan G. Woolpert
Director
Operations and Business

* COMMITTEE CHAIRMAN

STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from the Situational Variables and Efficiency of Concept Learning Project in Program 1. General objectives of the Program are to generate new knowledge about concept learning and cognitive skills, to synthesize existing knowledge, and to develop educational materials suggested by the prior activities. Contributing to these Program objectives, the Concept Learning Project has the following five objectives: to identify the conditions that facilitate concept learning in the school setting and to describe their management, to develop and validate a schema for evaluating the student's level of concept understanding, to develop and validate a model of cognitive processes in concept learning, to generate knowledge concerning the semantic components of concept learning, and to identify conditions associated with motivation for school learning and to describe their management.

ACKNOWLEDGMENTS

Since I cannot, after the course of more than five years, recall the names of all who gave scholarly, logistic, or moral support to the present study, I must content myself with the following inadequate list of those whose assistance was most critical in its completion:

Professor Julian C. Stanley, who suggested the topic, and offered continual encouragement and advice;

Drs. Ann Cleary, Margaret Harris, Chester Harris, Robin Fleming, Robert Hogan, and Peter Smith, who provided essential methodological and substantive suggestions;

Mr. James Bavry, who guided much of the data through the University of Wisconsin Computer Center;

"Mike" Patterson, Harriet Clutterbuck, Sue Johnson, Diane Brown, Deborah Stewart, and many others on the staff of the Wisconsin Research and Development Center for Cognitive Learning, who produced and reproduced the research materials and the report you are reading;

Professor Robert Davidson, my advisor, sponsor of this research; and

Camille Esselle Houston, who endured much while typing.

To these, and to the anonymous subjects, instructors, and homophone-finders who contributed time and effort, belongs the credit for whatever success this study has attained.

TABLE OF CONTENTS

Acknowledgments iv

List of Tables and Figures vi

Abstract vii

CHAPTER

I Introduction 1

II A List of English Homophones 5

III Research Design and Hypothetical Models. 9

IV Selection and Development of Instruments 22

V A Correlational Study of Homophone Recognition . . 34
- Instruments 34
- Subjects. 42
- Administration. 43
- Scoring 44
- Preliminary Results 45
- Summary of Preliminary Findings 58

VI Component and Factor Analysis of the Data. 59
- Determination of Comparable Factors 59
- Description and Interpretation of Factors . . . 64
- Sex Differences 73
- Factors of Experimental Tests 79
- Summary and Implications for Further Research 83

References 86

Appendix A: Homophone List. 89

Appendix B: Pilot Study Instruments 163

Appendix C: Experimental Instruments. 165

LIST OF TABLES

TABLE PAGE

4.1 Frequencies of Errors by Items and by Instructions for Homophone-Homonym Word Lists (N = 16). 26

5.1 Test Battery for Homophone Factor Study 35

5.2 Factorial Structure of Homonym Test. 40

5.3 Means and Standard Deviations by Sex, and Reliabilities of Variables 46

5.4 Correlations of Raw and Derived Homophone and Homonym Scores 52

5.5 Intercorrelations among Twenty-one Variables in the Homophone Study. 54

6.1 Factor Results for Homophone Matrices 81

6.2 Intercorrelations Among Oblique Factors 86

6.3 Uniqueness Estimates for Matrices A and B 93

6.4 Intercorrelations of Variables for Female and Male Subjects 95

6.5 Median Intercorrelations Among Nominal Clusters for Female and Male Subjects 98

LIST OF FIGURES

FIGURE PAGE

4.1 The Regression of TL Rank upon False-Negative Errors for Pilot Study Homophone-Homonym Word Lists (N = 16) 32

ABSTRACT

A homophone is a word having the same pronunciation as another English word, but a different spelling. A list of 7,300 English homophones was compiled, and used to construct two tests. Scores were obtained on these and on reference tests for J. P. Guilford's factors CMU, CSU, DMU, and DSU for 70 native speakers of midwestern American English from a university population. The homophone tests showed Hoyt reliabilities of .95 and .87 for these subjects.

Following Harris's procedure for determining Comparable Common Factors, the 15 x 15 matrix of intercorrelations was subjected to three factoring procedures, each yielding oblique and orthogonal solutions. Results were in close agreement for all analyses, yielding three common factors. Two corresponded to CMU and to DMU. The CSU and DSU tests loaded on the third factor, which had its largest loadings on the homophone tests, and involved comparing verbal stimuli with formal elements of internally generated lists.

These findings replicate Harris's failure to extract distinct CSU and DSU factors, and suggest that homophone recognition tasks can provide short but reliable reference tests for the symbolic factor into which CSU and DSU coalesce.

To Camille, Njal, and Tara Gillian Criseyda

. . . they're saying, "Piece. We want a piece."

--Unidentified peace officer in
Richard Leacock's film Chiefs

Chapter I

INTRODUCTION

Natural languages are ambiguous in that an utterance rarely specifies a unique meaning. The literature of classical antiquity provides many famous examples, such as the verbal trick whereby Homer's polytropos Odysseus outwitted Polyphemos after the destruction of Troy (Homerus, 1919). Calling himself "oudeis" (="No-one"), Odysseus blinded and escaped from the giant, whose howls of rage were construed by the other Cyclopes to mean that no one was harming him. Historically, the priest of Apollo at Delphi hedged his predictions by the use of syntactically equivocal language ("Pyrrhum quidem Romanos vincere"), so that whatever the fortunes of Dr. Skinner's namesake, the oracle would appear to be valid. The literary styles of the tragedian Euripides and of the scientific poet, Lucretius, conspicuously feature paranomia (puns), which were unquestionably indulged in by more facetious ancients who esteemed such use of language.

While ambiguity may frequently be desirable in literature, divination, and oratory, it poses severe problems in logic, philosophy, and law. Hellenes began the awesome task of "clarifying" language by attaching unique meanings to certain words. The Hellenistic era witnessed the labors of the Church Fathers to isolate the meaning of revelation; despite schismatics and heresiarchs, they secured a consensus of dogma which endured for a millenium. (Elsewhere similar enterprises, not necessarily independent, attached unique interpretations to the writings of Mohammed, Confucius, and Lao Tse.)

It is convenient to distinguish two operators in the processing of ambiguous information: a convergent lexic process which recognizes one meaning in an ambiguous verbal sign, and a divergent vatic process of entertaining multiple meanings. The two are not wholly antagonistic; the medieval West, whose philosopy achieved a unanimity unapproached by subsequent eras, was the heyday of allegory, a semantically divergent literary form. Danté himself, whose Commedia celebrates the theocentric Universe, has left us an elaborate theory of "polysemantic" meaning (1887).

A modern analogy exists with the logical operators AND and OR. If a sign admits a vector of possible meanings, the vatic process associates all of them with the signal, even if they are mutually contradictory:

(1) $V = S_1 \text{ OR } S_2 \text{ OR} \ldots S_n$,

for S_1, S_2, ..., S_n the possible meanings. The lexic process excludes, using some logical or external criterion, those meanings which are inconsistent with the context of the sign:

(2) $L = V \text{ AND } C$,

for C the contextual criterion. By choosing C so that $L = S_1$, (2) can always be arranged to yield an unambiguous meaning.

While formula (2) suggests the antecedence of the vatic process, note that one could just as well have written:

(3) $L=(S_1 \text{ AND } C) \text{ OR } (S_2 \text{ AND } C) \ldots \text{ OR } (S_n \text{ AND } C)$.

If in fact these hypothetical processes have any correspondence to actual cognitive operations, their temporal and developmental sequence should be the subject of empirical investigation, rather than of rational speculation, and they are offered here as tentative descriptive terms, rather than as explanatory mechanisms or members of an ordered hierarchy. The inability to perform one or the other type of interpretation may be detrimental in certain situations.

Note that V=L if C excludes nothing. What if C exdudes everything? Properly speaking, the interpreter could accept the sign as meaningless, just as "exdudes" in the previous sentence is not a standard word, and has no lexical meaning. Many readers, of course, will have perceived "excludes," the word demanded by the context, subconsciously or consciously dismissing the actual sign as an error. Thus C acts not merely as a mask to exclude inappropriate possible meanings, but it orders the series S_1, S_2, ..., S_n, and provides initial terms which may take priority over the usual meanings of a sign.

Thus, an important principle in textual criticism is lectio difficilior, selection of the more difficult reading, since low frequency words tend to be perceived as more commen words by manuscript copyists. An admirable example of Freudian parapraxis, and of a situation where lexic interpretation is less appropriate than vatic, occurs in one of Vladimir Nabokov's (1947) novels, where a beleaguered statesman delivers a speech comparing himself to "an animal who has lost his feet in a rising sea." The reader should take this to mean (S_1) that he is like a beast stumbling in the surf, who must struggle to escape drowning; (S_2) that he is like an animal evolving backwards, regressing into more primitive forms in a nightmare of helplessness; (S_3) that he is like a man too distraught to read correctly the text of his statement, which doubtlessly contained the more pedestrian figure of "an admiral who has lost his fleet."

The present research examines ambiguity whose locus is in individual words. Specifically, it is concerned with homophones, words which sound the same as other words. (These should not be confused with homonyms, words having the same written form as other words. Thus AGAPE (= yawning) and AGAPE (= love feast) are homonyms, whereas SERIES (= function of integers) and CERES (= Persephone's mother) are homophones.) Fowler (1965) uses homophone in the restricted sense of a word having the same sound as another word, but a different spelling. This is the sense in which the term will be used below.

The English language, with its huge vocabulary, its non-phonetic, etymologically conservative spelling, and its diverse sources is quite rich in homophones, so that native speakers necessarily acquire some skill in the interpretation of such words. This study investigates the processes by which homophones are recognized, and attempts to relate these to other verbal and cognitive operations.

Chapter II

A LIST OF ENGLISH HOMOPHONES

The intentions of the present study were basically psychometric: to relate the traits involved in homophone recognition to previously identified cognitive factors. Before data could be gathered, however, it was necessary to obtain some homophones. A list of all the English homophones is desirable for a number of reasons. It would provide a pool of items for subsequent research, and would enable test constructors to identify nonhomophones for use as distractors. Instruments constructed by randomly sampling items from such a list would permit generalization to the universe of English homophones.

To collect all of the English homophones, however, appears to be a chimerical quest. What is an "English word," and where is a list of such entities, that the homophone subset might be isolated? While one may be tempted to define "English words" as entries in some dictionary, such Rhadamanthian tautologies provide only subsets of the union of all English vocabularies. Lexicographers are no more infallible than educational psychologists, and it is absurd to imagine that a "comprehensive" dictionary includes all of the orthographic and phonemic forms which the potential respondents to a homophone test might regard as standard. Technical, regional, obscene, obsolete, novel, and other special forms which are intentionally or inadvertently omitted from reference works may be part of the active vocabulary of individuals, and to regard such entities as non-words can be psychometrically justified only if respondents are advised of these exclusions. The most easily communicated instruction is that a respondent regard two words as homphones if he pronounces them identically but spells them differently. Though no master list of homophones so defined can be constructed, the problems of valid keying are no more severe here than with items involving verbal or numerical analogies.

No attempt was made to secure an encyclopedic list of homophones. A standard reference work, Webster's Seventh New Collegiate Dictionary (hereafter WCD)[1] provided an initial list of 130,000 words, which the investigator inspected for homophones, using the same reference as an authority on pronunciation. While a more comprehensive dictionary would have contained more homophones, it was felt that these would have been chiefly low-frequency words, of little value as items, and that high-frequency homophones might more easily elude an investigator inspecting a longer list. The pool of homophones which were obtained in this way was augmented (by about 5%) with homophones from various other sources: homophone dictionaries, words collected by members of a test construction seminar, and words casually encountered by the investigator and his acquaintances.

The following criteria were generally followed in accepting or rejecting a possible homophone. If *any* lexically listed pronunciation of a word was the same as that of an orthographically different word, both words were included as homophones. Thus CHANCE and CHANTS, which WDC regards as homophones, are included, although other dictionaries[2] distinquish between their pronunciations. Some words (GRAY, GREY) have more than one acceptable spelling; if these are each appropriate to all themeanings of the word, then such forms were not included in the pool (unless some third orthographic form had the same pronunciation). But if a spelling variant is used with some but not all meanings of a word (INDICT, INDITE), then the spelling variants were considered to be homophones of one another.

---

[1] _______, Webster's Seventh New Collegiate Dictionary. Springfield, Massachusetts: G. & C. Merriam Company, 1963.

If two or more homophones are the same part of speech, inflected forms of those words (-s, -ed, -ing, -er, -est, -eth) may also be homophones. Sets of homophones involving an uninflected form are called primary sets; sets involving only inflected forms are called secondary sets. The homophone pool obtained by the procedure described above contained about 1700 primary sets, and more than 1600 secondary, each containing from two to nine homophones. The four largest sets were the following:

(2.1)

AIR
AIRE
ARE
AYRE
E'ER
ERE
ERR
EYRE
HEIR

CEES
C'S
SEAS
SEES
SEISE
SEIZE
SIS

CEROUS
CIRROUS
CIRRUS
SCIRRHOUS
SCIRRHUS
SEERESS
SEROUS

CREWS
CROUSE
CROUSE
CRUISE
CRUS
CRUSE
KREW(E)S

---

[2] E.g., C. L. Barnhart (Ed.), The American College Dictionary. New York: Random House, 1960.

Appendix A displays the sets of words which the present investigator classified as homophones. Proper nouns and proper adjectives generally were not included. This list is intended to aid in the construction of homophone instruments. Persons using it are urged to refer to the dictionaries of their choice, in order to acquaint themselves with any variant pronunciations. Many of the words included are not homophones according to other authorities; but any word which appears on this list is probably unsuitable for use as a distractor in a test of homophone recognition. Sets of homophones omitted from this list can be presumed to involve chiefly technical or low-frequency words, spellings, or pronunciations, but no claim of completeness with respect to high-frequency words is offered. To the knowledge of the present investigator, this is the largest collection of English homophones ever assembled.

Chapter III

RESEARCH DESIGN AND HYPOTHETICAL MODELS

The research strategy for the present investigation was fairly simple. One or more instruments would be developed yielding scores on a homophone task; the covariance structure of these scores and of selected measures of cognitive abilities would be examined by factor analytic or similar techniques, in order to describe performance on the homophone tests in terms of previously identified cognitive traits.

While the orientation of this study was primarily empirical, some rational basis was felt to be necessary for the initial selection of homophone items, for the selection of respondents from whom the measures would be obtained, and for the choice of marker tests to be intercorrelated with the homophone scores.

Practical considerations suggested the choice of "recognition" items, whereby subjects would be presented with written lists of words, and be instructed to indicate those words which were elements of homophone sets. It was anticipated that respondents would be available only for relatively brief (one- to two-hour) periods of time, during which this format would permit the presentation of large numbers of items. Scoring would also be simpler, under these procedures, than with constructed responses, and items which correlated poorly with total test scores could be excluded post hoc without seriously restricting the range of scores. The use of paper and pencil items would additionally facilitate the collection of data by permitting the group administration of tests.

The advantantages of this item format are gained at some cost. The use, for example, of aural stimuli, while introducing problems of misperception, would have avoided the arbitrary selection of one element of a homophone set as stimulus. Constructed responses would have provided information on the order in which set elements were recalled, and on the types of errors involved in false positive responses. The use of more elaborate apparatus, or of individual testing, could have provided response latencies, a potentially more sensitive source of dependent variables than simple right-wrong scores. It was felt, however, that such refinements could await further research efforts in this area.

Logistic considerations were also important in the decision to use a university population as the source of subjects; members of Educational Psychology courses at a large midwestern university constituted the most accessible pool of potential respondents, whose participation in such research efforts was encouraged by departmental requirement. The comparatively high level of verbal ability typical of university students would facilitate test construction, by permitting the use of medium-frequency words without flooding the data with variance attributable to ignorance of word meanings or of orthography. While it would be interesting to study the development of homophone abilities in the lower grades, the number of items suitable for respondents with less knowledge of vocabulary and spelling is correspondingly smaller.

More elaborate speculation guided the selection of marker tests for use in the subsequent analysis. The following attempt was made to provide models of the covert processes involved in classifiying a stimulus word as a homophone or non-homophone:

Assume that three modes are available for the internal representation of verbal stimuli: aural, graphic, and semantic. Let A, G, and S symbolize the representation of a stimulus in each respective mode.

Assume that there exist associative operators which transform information from one mode to another. These operators will be symbolized by solid right-pointing arrows, and be specified by the modes which they link.

Thus

(3.1) $G \rightarrow A$

associates a sound with a written form, while

(3.2) $A \rightarrow G$

performs the inverse of (3.1). The operator

(3.3) $G \rightarrow S$

assigns a meaning to a written form, while

(3.4) $A \rightarrow S$

attaches a meaning to a sound. One could define simple inverses to (3.3) and (3.4), giving a symmetrical set of six binary operators; but the elegance of this formulation must be weighed against verisimilitude.

In completing crossword puzzles or consulting a thesaurus, one may attempt operations of the forms

(3.5) $S \rightarrow G$ or

(3.6) $S \rightarrow A$

but the time required is vastly longer than for operators (3.1-4). Under the circumstances of interest to the present study, semantic-mode information is initially available in aural or graphic form, so that

(3.7) $(A \rightarrow S) \rightarrow G$ and

(3.8) $(G \rightarrow S) \rightarrow A$

describe the triple-argumented operators which rapidly associate semantic mode information with graphic or aural forms. For simplicity, the parentheses in (3.7,8) will not be written, since arrows proceeding from an S will always be intended as (3.7) or (3.8) in the discussion below, and never intended as (3.5,6). An example of (3.8) is: TEAR → (= "tear-drop") → /tir/, or TEAR → (= "rip") → /tār/. (Throughout this paper, graphemic representations will be written in upper case letters; semantic representations will be written in parentheses and quotes, preceded by an equality sign; and aural representations will be written phonetically and between slashes.)

It is convenient to further ornament this notational system with another operator, the identity test

(3.9) $(Q_i - Q_j)\begin{array}{c}\dashrightarrow \\ \dashrightarrow\end{array} .$

Here $Q = A$, G, or S, and Q_i and Q_j are potentially different representations of information in the same mode. The identity test provides a yes or no decision to the question, "Is Q_i the same as Q_j?", and the broken arrows lead to the operation which follows each outcome (upper if positive, lower if negative). Using this notation, it is possible to specify various models of the processes involved in interpreting homophones and homonyms.

The following model describes a strategy for classifying visual stimulus X as a homophone. The symbol Y represents some procedure to give out an overt response that X is indeed a homophone; the double arrow following X represents some procedure for coding X as G.

(3.10)
$$\begin{tikzcd} X \arrow[r, Rightarrow] & G_o \arrow[r] & A \arrow[r] & S \arrow[r] & (G - G_o) \arrow[r, dashed] \arrow[lll, dashed, bend right] & Y \end{tikzcd}$$

In words, a word is read, and its sound is (vocally or subvocally) obtained, as is its meaning. (Whether S in fact proceeds from G, from A, or from both need not concern us here.) The graphemic form of the word with that sound and meaning is compared with the original written stimulus, (G_o); if these are identical, the aural representation is reconsidered, and a meaning is found again. The spelling of this semantic form is compared with the original stimulus, and the process is repeated until an orthographic discrepancy signals the discovery of a homophone. Note that no provision is made here, nor in the other models given below, for the abandonment of the search for homophones. Presumably some time intercept permits the subject to go on to the next item when no homophone can be found.

Note also that (3.10) is a fallible strategy. If $G_o \rightarrow A$ can yield more than one pronunciation, the initial choice of an A having no homophones may lead to a false negative. For example, BASS → /bās/ → (="fish") has no homophones, whereas BASS → /bās/ → (="musical term") has the homophone BASE. A less fallible model, then, is

(3.11)
$$\begin{tikzcd} X \arrow[r, Rightarrow] & G_o \arrow[r] & A \arrow[r] & S \arrow[r] & (G - G_o)' \arrow[r, dashed] \arrow[lll, dashed, bend right] & Y \end{tikzcd}$$

though there is no guarantee that the respondent will ever choose the critical pronunciation.

An alternative strategy is

(3.12)
$$\begin{tikzcd} X \arrow[r, Rightarrow] & G \arrow[r] & A \arrow[r] & (S - S_o)' \arrow[r, dashed] \arrow[l, dashed, bend right] & Y \end{tikzcd}$$

which saves a step, and which will correctly solve certain items. This yields false positives, however, in cases of isophonic homonyms (words such as FALL (="autumn") and FALL (="plunge") having one pronunciation, but distinct meanings). A variant of (3.12),

(3.13)
$$\begin{tikzcd} X \arrow[r, Rightarrow] & G \arrow[r] & A \arrow[r] &)S - S_o)' \arrow[r, dashed] \arrow[ll, dashed, bend right] & Y \end{tikzcd}$$

will produce false positives for allophonic homonyms (words such as MINUTE (="sixty seconds") and MINUTE (="tiny") having distinct pronunciations and meanings), as well as isophonic homonyms.

A different model hypothesizes the graphemic modification of the stimulus:

(3.14)
$$\begin{tikzcd} X \arrow[r, Rightarrow] & G_o \arrow[r] & A_o \arrow[r] & (G-G_o)' \arrow[r, leftrightarrow] \arrow[l, dashed, bend right] & S \arrow[r] & (A-A_o)' \arrow[r, dashed] \arrow[llll, dashed, bend left] & Y \end{tikzcd}$$

Here a stimulus (such as DJINN) would be pronounced and subjected to orthographic transformations which tend to preserve the aural form (DJIN, JIN, JINN, JINNH, GINNH, GINN, GYNN, GYN, GIN, etc.). If a meaning can be assigned to any of these graphic forms, its pronunciation is compared with the original stimulus; if identical, a homophone has been found. To guard against false positives in the case of orthographic variants of the same semantic form (such as DJINN, JINN, and JINNH, (="plural of JINNI")), another identity test could be included:

(3.15) $$\begin{tikzcd} X \arrow[r, Rightarrow] & G_o \arrow[r] & S_o \arrow[r] & A_o \arrow[r] & (G-G_o) \arrow[r] \arrow[loop above, dashed] & (S-S_o) \arrow[r] \arrow[l, dashed, bend right] & (A-A_o) \arrow[r, dashed] \arrow[ll, dashed, bend left] & Y \end{tikzcd}$$

Much simpler, more fallible strategies such as

(3.16) $$\begin{tikzcd} X \arrow[r, Rightarrow] & G_o \arrow[r] & A \arrow[r] & (G-G_o) \arrow[r] \arrow[loop above, dashed] & S \arrow[r, dashed] & Y \end{tikzcd}$$

are available; these would give false positives if the spelling change resulted in a differently pronounced word (INCITE, INSIGHT), or an orthographic variant.

Formulae (3.10-16) offer a profusion of models; these are proposed not with a view towards validating one or more of them, but in order to suggest cognitive factors which may be involved in homophone recognition. The discussion which follows borrows the terminology of J. P. Guilford's "Structure of Intellect" taxonomy of cognitive factors. Guilford's well-known model hypothesizes no fewer than 120 distinct mental abilities, arranged in a three-dimensional rectangular lattice, factor names being provided by the three coordinates within this lattice.

One dimension of Guilford's model is "operation," which comprises five categories: Cognition, Memory, Divergent Production, Convergent Production, and Evaluation. Taking these in reverse order, "Evaluation" is defined as

> ...a process of comparing a product of information with known information according to logical criteria...[1]

---

1. J. P. Guilford, The Nature of Human Intelligence, McGraw-Hill, New York, 1967, p. 185.

While this appears to describe the identity tests of (3.10-16), it was decided not to investigate the involvement of evaluative factors. The grounds for this were that tests of synonym recognition load on the cognitive, rather than evaluative, plane of Guilford's lattice; these seem to correspond to $(S-S_o)$ in (3.12, 13, 15). Beyond this, it was felt that the $(G-G_o)$ and $(A-A_o)$ tests in (3.10, 11, 14-16) might contribute variance to the data not because of differential abilities to perform such tests, but according to whether respondents selected a strategy which included or failed to include such a test. Performance on evaluative tasks has no *a priori* reason to predict strategy selection, so such tasks could be expected to have little explanatory value.

"Convergent Production" is described as

> ...the prevailing function when the input information is sufficient to determine a unique answer. [2]

This closely corresponds to (3.7, 8); again, however, the operator was considered unimportant in the present task, since the anticipated variation in the ability of respondents to assign meanings to verbal stimuli would be related to vocabulary size, whose association with the "Cognitive" operator is well established.

"Divergent Production" is identified with "fluency, flexibility, originality," "elaboration," and "verbal versatility."[3]

2. J. P. Guilford, *op. cit.*, p. 171.

3. *Ibid.*, p. 138.

This was expected to be a critical operation in homophone recognition; it appears in (3.10-16) as intersecting arrows, where G→A, A→S, or A→G are iterated after equality tests. According to these models, homophone recognition is impossible unless at least one of these operators produces non-unique representations of the stimulus. Accordingly, several divergent-production factors were chosen for inclusion in the final battery.

"Memory" is defined by Guilford as

> ...retention or storage, with some degree of availability, of information in the same form in which it was committed to storage, and in connection with the same cues with which it was learned. [4]

It was anticipated that memory would play a very minor role in the homophone task; while the models require that the stimulus and its transformations be retained, whatever information might be forgotten could quickly be reconstructed by reference to the printed test materials. The retrieval of vocabulary and spelling information comes under the "Cognitive" rubric, rather than under "Memory"; the latter operation was neglected in the present study.

"Cognition" is the most primitive of Guilford's operations, a necessary condition for the others to occur. It is defined as

> ...awareness, immediate discovery or rediscovery, or recognition of information in various forms... [5]

---

4. *Ibid.*, p. 211.

5. *Ibid.*, p. 203.

This corresponds to the operators defined in (3.1-8), and would seem to be involved in any model of homophone recognition. Several cognition factors were chosen as being likely to have explanatory value in the present study.

These were the five categories of Guilford's "Operation" dimension. A second dimension, "Products," consists of six somewhat hierarchic categories (Units, Classes, Relations, Systems, Transformations, and Implications) having reference to the complexity of the outcome of an "Operation." Since the homophone task restricts itself to the properties of isolated words, the simplest of these categories, "Units," appeared to be the only relevant one. All of the factors for which marker tests were selected involved units.

The remaining dimension of Guilford's model is "Content," comprising four categories: Figural, Symbolic, Semantic, and Behavioral. These

> ...are the very broad, substantive areas of information, whereas the six intersecting product categories are formal types of differentiation. [6]

This represents a fairly radical departure from the tendency of earlier theorists, such as Burt, Vernon, or Cattell, to dichotomize this domain.

The "figural" category denotes information

> ...in concrete form, as perceived or as recalled in the form of images. [7]

---

[6] Ibid., p. 221.

[7] Ibid., p. 227.

While this may seem to describe the G and A modes of the models proposed in (3.10-16), these appear to be pre-empted by the "Symbolic" category, which includes information

> ...in the form of signs, materials,
> ...letters, numbers, musical notation,
> and other "code" elements. [8]

It was decided to include tests loaded on symbolic factors in the final battery, and to neglect figural factors.

The "Semantic" category of content in Guilford's model is wider than the Semantic mode postulated above in this chapter. Though he states that

> Semantic information is in the form
> of meanings to which words commonly
> become attached, [9]

Guilford amplifies this with nine additional pages of discussion wherein he includes connotative and non-verbal meaning under the semantic rubric. It was anticipated that semantic factors would be of considerable importance in accounting for the variance of any homophone data.

The last category, "Behavioral" information, is defined as

> information, essentially non-verbal,
> involved in human interactions, where
> awareness of attention, perception,
> thoughts, desires...of other persons
> and of ourselves is important. [10]

---

8. Ibid., p. 227.

9. Ibid., p. 227.

10. Ibid., p. 77.

It appears to have no relevance to the present study.

On the basis of the rational considerations indicated above, four of Guilford's factors appeared to be most likely to be involved in homophone recognition: cognition of semantic units (CMU), divergent production of semantic units (DMU), cognition of symbolic units (CSU), and divergent production of symbolic units (DSU). The first, CMU, appears in the models (3.10-16) as G→S and A→S, and is empirically related to vocabulary knowledge. The second, DMU, appears in the models as iterative G→S and A→S. The next, CSU, appears in the models as non-iterated G→A, A→G, S→G, and S→A; it has elsewhere been associated with knowledge of spelling (Hoepfner, Guilford, & Merryfield, 1964). Guilford in fact distinguishes two factors which occupy the CSU cell in his model, a visual (CSU-V) and an auditory (CSU-A) factor. The latter may be related to Carroll's "phonetic coding ability " (1962) and is of interest in models (3.14-16), but is less well established by previous studies, and is marked by tests involving auditory stimuli; the relation between encoding external stimuli, and encoding internally stored or generated information, was considered likely to be quite tenuous, and to involve irrelevant perceptual abilities in the former case. Only CSU-V was examined. The last chosen factor, DSU, appears in the models as iterative G→A and A → G, and is of great importance in (3.14-16).

Factors CMU, DMU, and DSU are fairly "robust," and are each classified by Guilford and by French as having been detected in ten or more empirical studies. They correspond closely to elements of other theories of cognitive factors: DMU = French's Ideational Fluency (Fi) = Cattell's "Universal Index T6"; DSU = Word Fluency (Fw) = U.I. T15; CMU = Verbal Comprehension (V) = U.I. T13 (French, Ekstrom, & Price, 1963). The fourth factor, CSU-V, is less well established; but since tests which mark it tend also to be loaded on CFU (identified with French's Speed of Closure and Cattell's U.I. T3), it was decided to include this factor on the chance that its inclusion would be informative. At the time the present study was being planned, preliminary results of Harris's (1969) re-analyses of some of Guilford's data showed that CSU coalesced with DSU, an interesting finding which the present study could attempt to replicate.

Chapter IV

SELECTION AND DEVELOPMENT OF INSTRUMENTS

As a preliminary step toward the development of a test of homophones, a pilot study was conducted with a view towards: (1) identifying test items of appropriate difficulty; (2) determining suitable time limits for tests using such items; and (3) generally exploring the domain of homophone recognition.

Two instruments were prepared for the pilot study. One was a list of 48 words, having the following factorial structure (in the Fisherian, rather than Thurstonian sense): (1) half of the words were members of homophone sets, according to the list described in Chapter II above, and half were not; (2) two-thirds of the words were members of sets of homonyms, and one-third had only one definition listed in WCD; (3) one-third of the words were members of sets of allophonic homonyms (pairs of words having the same spelling, but different pronunciations and meanings). Items were arranged in eight blocks of six, each block consisting of elements of the form

(4.1) HNA,
HNa,
Hna,
hNA,
hNa,
hna.

Here H = homophone, N = homonym, A = allophonic homonym; corresponding lower case letters indicate membership in the converse set. These six elements were arranged in a random order within each block. Note that a complete orthogonal design is here impossible, since A is a subset of N; HnA and hnA cannot occur.

Items satisfying the restrictions of this design were selected non-randomly from the homophone list and from WCD. It seemed likely that the use of randomly selected items would result in a test heavily loaded on CMU, owing to the preponderance of low frequency words. In addition, rather few words of the forms HNA, Hna, and hNA exist, so that filling these cells by sampling from some general list of words would be a tedious undertaking of marginal value. In filling the homonym cells (HNA, HNa, Hna), the investigator selected as stimuli relatively common words, whose homophones tended to be common words. The Thorndike-Lorge (1938; hereafter TL) list of 40,000 words was used as a guide to word frequencies. The abscissa in Figure 4.1 below shows the TL frequencies of the highest frequency homophones of the 24 positively keyed homophone items. (These frequencies are approximate ordinal positions among the words included in the TL count; thus a word with a TL frequency of 1750 is probably more common than a word with a TL frequency of 8500.) Some lower frequency words were included in order to examine the regression of TL frequency upon item difficulty.

The motivation for including I and A items deserves some comment. It was hypothesized that respondents who used strategies equivalent to (3.13) would report false positives for stimuli of the form hIA or hIa; (3.12) would lead to false positives for hIa stimuli. Respondents using (3.10,12, 14-16) would tend to report false negatives for HIA items, particularly if the more common homograph had homophones.

The second instrument used in the pilot study was a selection of verse, comprising twenty lines of iambic pentameter, or ten heroic couplets. This material was based chiefly upon various portions of Alexander Pope's Essay on Man, and included some eighty words from the list of homophones. The instrument was constructed by reading through Pope and selecting couplets which contained two or more homophones. Couplets were occasionally modified by the investigator, and were sequenced in such a way as to maintain some semblance of logical continuity. The purpose of this second instrument was twofold: to examine additional homophone items, and to investigate the effect of "context" upon the recognition task. The metrical format was chosen because of a belief that any prose passages which might be contrived to contain a high proportion of homophones would be conspicuous as such; stilted diction would serve as cues to respondents, so that the instrument would tend to measure stylistic sensitivity, an interesting but irrelevant trait. Since a high homophone density was necessary, it was hoped that an attempt at Neo-classic verse would neutralize respondents' capacity to detect unnatural language by overloading it. In addition, the sonorous structure provided by a rhyme and meter was expected to mask the semantic discontinuity of the passage. Appendix B exhibits both of the instruments used in the pilot study.

Subjects of the pilot study were sixteen students at the University of Wisconsin. Four were undergraduate students; three were men and thirteen were women. About half were acquaintances of the investigator, and the rest were enrollees in an Educational Psychology course, who cooperated in order to satisfy departmental requirements of experimental participation. Ages ranged from 19 to 25 years. While this was a very small and non-random group, it seemed adequate for the exploratory purposes of the pilot study.

These instruments were administered individually to each subject by the investigator, who read the instructions and remained present to record the time spent on each test. Instructions were revised until they reached the form displayed in Appendix C. One of the most persistent sources of confusion was the fact that pairs of homophones were not present in the test materials. Subjects recorded their responses on the test booklets, circling items which they classified as homophones or homonyms. The first instrument, the word list described in (4.1), was administered three times in succession to each subject, with instructions to indicate homophones, homonyms, and allophonic homonyms, always in that order. Besides the three copies of the first instrument, test booklets contained two copies of the second, separated by carbon paper. After four minutes of work on the second instrument, subjects were instructed to turn to the second copy and to continue underlining homophones. In this way, scores for four minutes and for unrestricted time allotments were available.

Results for the first instrument are summarized in Table 4.1. In general, most items showed little variability, and distractors elicited few false positive responses for all instruction conditions.

Table 4.1

Frequencies of Errors by Items and by Instructions for Homophone-Homonym Word Lists (N = 16)

		Instructions					Instructions		
Code	Item	H	N	A	Code	Item	H	N	A
11	not	0	2	0	41	miss	0	3	0
12	pail	0	0	0	42	flatter	0	7	0
13	ale	3	0	0	43	well	1	0	0
14	boar	2	1	0	44	spoke	5	3	0
15	quartz	8	0	0	45	cross	0	0	0
16	boulder	7	1	0	46	spring	0	0	0
17	murderess	16	0	0	47	general	0	2	0
18	grocer	16	0	0	48	hail	1	11	0
	TOTAL Hna:	42	4	0		TOTAL hNa:	7	26	0
21	remind	0	0	0	51	desert	14	11	11
22	onto	0	0	2	52	lead	12	5	4
23	quickly	0	0	0	53	tears	13	2	2
24	obstacle	0	1	0	54	does	16	15	12
25	detests	0	0	0	55	bow	12	0	5
26	cactus	0	0	0	56	bass	9	2	4
27	opal	0	2	0	57	wind	13	1	2
28	discuss	0	0	0	58	muse	15	7	15
	TOTAL hna:	0	3	2		TOTAL HNA:	104	43	55
31	might	9	9	0	61	wound	1	6	3
32	plain	0	9	0	62	invalid	1	2	2
33	peer	6	0	0	63	minute	0	3	2
34	brood	15	5	0	64	dove	0	6	4
35	rung	13	0	0	65	content	0	7	2
36	laps	14	1	0	66	unionized	0	15	16
37	band	15	0	0	67	resumes	0	9	11
38	lean	14	0	0	68	number	0	12	14
	TOTAL HNa:	86	24	0		TOTAL hNA:	2	60	54

H = Identify homophones.
N = Identify homonyms.
A = Identify allophonic homonyms.

Under the instruction to indicate homophones (H), the order of increasing difficulty among positively keyed item categories was Hna (3.19 errors per subject), HNa (5.38), and HNA (6.50). Since 8.00 is the maximum number of errors, items of the last two categories tended to be quite difficult; six subjects missed all of the HNA items. Among negatively keyed item categories, the order of increasing difficulty was hna (0.00), hNA (0.13), and hNa (0.44). Of the nine false positive errors in the data, five were in response to the hNa item SPOKE.

Interpretation of these results is complicated by the non-orthogonality of the categorical variables to TL frequency. Among positvely keyed items, the HNA and Hna categories had homophones with median TL values of 2750 each, while the HNa value was 7500. Thus while models (3.12, 13) predict the HNa < Hna with respect to difficulty, the observed reversal of this relationship may be due to the less common words involved in the HNa items. The observed Hna < HNA also contradicts (3.13), despite the approximate equality of word frequencies; this result is consistent, however, with models (3.10, 12, 14, 15, and 16), which assume the uniqueness of $G \rightarrow A$. The observed HNa < HNA tends to support these latter models, and to discredit (3.11).

The data for the negatively keyed items are equivocal. The observed relationships hna < hNA, and hna < hNa, are predicted by (3.12, 13), and the success of SPOKE as a distractor is difficult to account for under the other models. In every instance, homonyms incorrectly classified as homophones were correctly identified as homonyms when the same subjects were asked to do so under the next set of instructions. On the other hand, the very low rate of false positives is inconsistent with models so fallible as (3.12, 13); the hNa< hNA implied by (3.13) was not observed; and it is not inconceivable that five university students believed that /spōk/ → (="did speak") and /spōk/ → (="wheel part") had different spellings, since many examples of non-standard orthography occurred in the constructed response items of the subsequent main study.

Under the homonym (N) instructions, somewhat similar results were obtained. Non-positively keyed items showed vanishingly small error rates: 0.25 for Hna, 0.19 for hna, out of a possible maximum of 8.00. Among the positively keyed items, the modal number of errors was zero, and the ordering of increasing difficulty by category was HNa (1.50), hNa (1.63), HNA (2.69), and hNA (3.75).

Models of homonym recognition have been proposed above in (3.12, 13); two others can easily be constructed by modifying (3.10) to read

(4.2) $$X \Rightarrow G_o \to A \to \underline{S \to (G - G_o)} \dashrightarrow Y,$$

(4.3) $$\begin{tikzcd} X \arrow[r, Rightarrow] & G_o \arrow[r] & A \arrow[r] & S \arrow[r] & (G - G_o) \arrow[r, dashed] \arrow[lll, dashed, bend left] & Y. \end{tikzcd}$$

Since (3.12, 13) implies a fairly high error rate, the data tend to support (4.2, 3). All four models predict hna < Hna, HNa < hNa, and HNA < hNA, the observed relationships with respect to difficulty. The considerable increase in difficulty between (HNa, hNa) and (HNA, hNA) is better accounted for by (3.12) or (4.2) than by (3.13), or (4.3).

The third set of instructions, to identify allophonic homonyms (a), could be accomplished by a strategy such as

$$\begin{tikzcd} X \arrow[r, Rightarrow] & G \arrow[r] & S \arrow[r] & (A - A_0) \arrow[r, dashed] \arrow[ll, dashed, bend right] & Y. \end{tikzcd} \tag{4.4}$$

Results were somewhat paradoxical. Error rates for items of the form HNA (3.44) and hNA (3.38) were almost identical. For negatively keyed items, categories Hna, HNa, and hNa displayed error rates of zero, while the hna item ONTO inexplicably elicited two false positive responses, perhaps from subjects who construed it as a prefix. Errors in the HNA and hNA categories tended to be correlated with errors on the same items under N instructions, but in many cases such words were missed under N but correctly identified under A. These latter instances tended to support the hypothesis that strategies which test the uniqueness of G → A are available (since false negative error rates under A are below 50%) but infrequently used under conditions H or N.

One would do well to turn from speculative inferences based upon these data (so small a body that no further statistical treatment seemed worthwhile) to the chief purpose of the pilot study, the identification of suitable items for a test of homophone recognition. It is well known that the reliability and validity of a test is maximized for a given group when all items are of 50% difficulty for that group. If five to eleven errors (28% to 72% error rates) is taken as a satisfactory approximation of 50% difficulty, then under instruction condition H, only 12½% of the 48 items fell within this range. For conditions N and A, the corresponding figures are 25% and 6¼% respectively, as can be derived from Table 4.1 above. Since most of the items used in the pilot study were of inappropriate difficulty for the group tested, it was of the utmost importance to characterize the near-50% difficulty items, so that other items like them might be included in subsequent versions of the instrument.

Neither TL rank nor the homophone-homonym classificatory variables seemed adequate to account for the observed variation in item difficulty. For example, item 13, ALE, had a 19% error rate; its homophone AIL has a TL rank of 5500. Item 15, QUARTZ, has a homophone QUARTS with a TL rank of 2250, and is, like 13, a Hna item; yet half of the subjects failed it. Item 33, PEER, offers another example; this and item 34, BROOD, are in the HNa category, and both have homophones (PIER, BREWED) with TL rank of 5500. Yet 37% of the subjects failed item 33, while 94% failed item 34; these data were collected within 100 miles of Milwaukee.

In the hope of discovering other variables of explanatory value, a bivariate plot of errors by TL rank was made of the 24 positively keyed homophone items. After considerable trial and error, it was concluded that the non-linear regression of the following variables upon item difficulty well fitted the data:

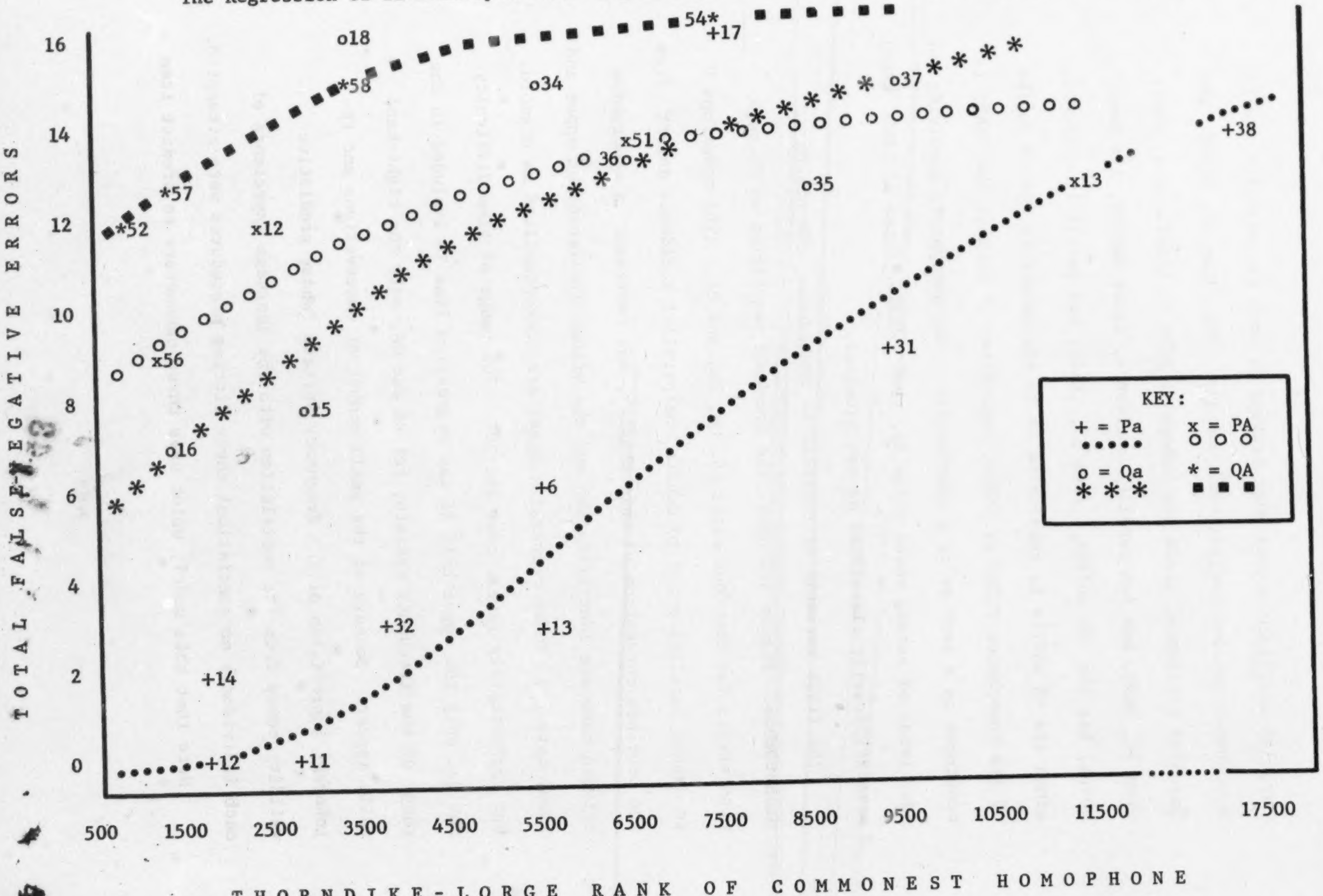
I. Frequency of homophone

II. Polyphonism (A versus a)

III. Parallel accidence (P, versus non-parallel accidence, Q)

The first two variables have been described above. "Non-parallel accidence" occurs when an item is a member of a primary set of homophones, as defined in Chapter II above, containing an inflected word.

Figure 4.1
The Regression of TL Rank Upon False-Negative Errors for Pilot Study Homophone Items (N=16)



Parallel accidence occurs when neither or both the stimulus and its homophones possess inflectional endings. Thus item 31, MIGHT, has parallel accidence, since its homophone MITE is uninflected; while item 37, BAND, has non-parallel accidence, since BANNED, its homophone, has the -ED suffix. Item 53, TEARS, has parallel accidence, since its -S suffix is considered to be equivalent to the -S suffix of its homophones TIERS or TARES, regardless of whether the TEARS is construed as a verb or as a substantive. The homophones involving the past tense of strong verbs (item 35, RUNG (WRUNG); item 52 LEAD (LED)) were arbitrarily classified as non-parallel.

The first variable is regarded as continuous, the others as dichotomous. Figure 4.1 shows the freehand regression of TL rank upon errors for the four cases Pa, Qa, PA, and QA. (The notations P to denote parallel and Q to denote non-parallel accidence depart from the previous convention of this chapter, but represent a compromise between mnemonic identification and the visual similarity of upper and lower case p.) These regression lines are conceptualized as sigmoid, but approximately linear over the 30% - 70% range of item difficulty. For Qa, only the upper half of the regression line is included in the range of the frequency variate; for PA and QA, only the right-hand tails appear. Because of the small number of observations and the inherent imprecision of the frequency variate (whose predictive validity comes from its correlation with the language experience of each individual) no statistical curve fitting procedures were attempted.

Note that this model, which uses three parameters to predict item difficulty, neglects N (the homonym - non-homonym status of the item) and the TL rank of the stimulus, as well as other potential predictors, such as the frequency of the graphemes wherein the homophones differ. No distinction is made between Q items in which the inflected or the uninflected form is the stimulus. The greatest deviation from the regression curve (item 17, MURDERESS[1]) is probably a miskeyed item, or would be regarded as such by many persons.

The second instrument, containing homophones in a verse context, appeared to be psychometrically superior to the first, though requiring more time to complete (about seven minutes, versus less than four for the first under each condition). Thirty-three items (41%) were in the 5 - 11 error range, and no item was answered correctly by all subjects. The context appeared to increase the difficulty of the task; the item NOT, which was passed by all subjects on the first instrument, was failed by 50% on the second; this was the only item appearing on both tests. Scores after four minutes were plotted against scores with time unrestricted; these showed a strong positive correlation. Scores appeared to reflect large individual differences: the number correct ranged from fourteen to fifty (both extremes scored by female graduate students in Educational Psychology), with a Kuder - Richardson Formula 21 reliability of .81 estimated from a mean of 30.4 and a variance of 94.6 (Cronbach 1960, p. 141). Only the word PRESENT elicited false positive responses, receiving two. Examination of subsets of the items proved inconclusive, perhaps an inevitable consequence of the small number of subjects observed in this pilot study.

[1] WCD regards MURDEROUS as a homophone of this word; The American College Dictionary does not.

Chapter V

A CORRELATIONAL STUDY OF HOMOPHONE RECOGNITION

Instruments

The pilot study described in the previous chapter and the theoretical considerations discussed above in Chapter III provided the basis for a larger scale investigation of homophone recognition. Arrangements were made to secure a group of subjects, and a battery of tests was selected which included experimental tests of homophone and homonym recognition, together with reference tests for factors DSU, DMU, CMU, and CSU. Table 5.1 lists the tests which composed this battery, Guilford's names for the factors associated with the reference tests, and the amount of time, exclusive of instructions, which was recommended for the completion of each test. The code names on the left are those used in the subsequent analysis to identify each variable.

Suffixes requires the respondent to write as many words as possible which end in a given suffix. Word Beginnings and Word Fluency are similar, the respective tasks being to provide words which begin with a given letter, and to provide words which include a given letter. All three are associated with Guilford's factor DSU, and involve the production of isolated words satisfying some phonetic restriction, without regard for meaning.

TABLE 5.1

TEST BATTERY FOR HOMOPHONE FACTOR STUDY

Code Name[a]	Title	Referent[b]	Items	Time[c]
ASF-DSU	Suffixes	DSU	1	4
BPR-DSU	Word Beginnings	DSU	2	6
CFL-DSU	Word Fluency	DSU	3	4
AIF-DMU	Ideational Fluency	DMU	2	6
DOC-DMU	Consequences[d]	DMU	3	6
CPT-DMU	Plot Titles[e]	DMU	2	6
AGZ-CMU	Verbal Comprehension	CMU	24	4
BVI-CMU CVJ-CMU	Advanced Vocabulary[f]	CMU	36	8
AOM-CSU	Omelet	CSU	30	4
BDV-CSU	Disemvowelled Words	CSU	25	5
BLISS	Bliss	-	78	4
HPHON NOPHO	Homophones[g]	-	48	4
HMNYM NONYM	Homonyms[g]	-	48	4
				65

[a] Code Name identifies variable in subsequent analyses.
[b] Referent is factor associated with test by Guilford.
[c] Time is in minutes. Column sum is 65.
[d] Test yields two scores; the second, Remote Consequences (6 RMCSQ), was also included in the analyses.
[e] Test yields two scores; second was discarded.
[f] Separate halves were each used as variates.
[g] Positively- and negatively-keyed items yielded separate scores.

Ideational Fluency requires the respondent to list as many objects as possible which lie in the intersection of certain descriptive categories. Because a total of only 90 minutes would be available for test administration, only the first two sections of this four-part test were used. *Consequences* requires that consequences of certain hypothetical situations be listed. Again, because of time limitations, only the first three of the five parts of this test were used. Scoring of this test distinguishes between "obvious" and "remote" consequences, so that the total number of each is separately reported. The distinction is between high and low frequency responses; the latter variable is taken as a measure of originality, rather than of fluency. *Plot Titles* asks the respondent to list appropriate titles for a brief story. As with the *Consequences* test, two scores are obtained, for high and low quality titles. Clever titles, like remote consequences, are associated with Guilford's factor DMT (divergent production of semantic transformations). Scores on *Ideational Fluency*, and the total number of obvious consequences and non-clever plot titles are supposedly measures of DMU, and involve the production of words and phrases satisfying some simple semantic restriction.

Verbal Comprehension (Form PX) is a vocabulary test from the Guilford-Zimmerman aptitude survey. The test has 24 items, on each of which a synonym must be selected from among five options. The *Advanced Vocabulary* test is a multiple choice vocabulary test in the same five-option format, consisting of two separately timed 18-item parts. For reasons indicated below, both halves were used as variates in the factor and component analyses. Both tests generally load on CMU, and involve recognition of synonyms.

The Omelet test presents four-letter anagrams; the respondent indicates the ordinal position of the initial letter of the solution. The test has two separately timed 15-item parts. Another test developed by Guilford and his associates, Disemvowelled Words, requires that the subject supply the missing vowels of 25 words. Some allowance for misspelling is made. As with the Omelet and vocabulary tests, the score is the total number of items correct. Both tests are designated by Guilford as markers of CSU-V, and involve the construction of words from given sets of alphabetic characters.

In addition to these ten standard instruments, three experimental tests involving homophones and homonyms were included in the battery. These were revisions of the two instruments described in the previous chapter.

The first of these, Homophone, was a 48-item revision of the list of words used in the pilot study. It was revised from the original list according to the following principles:

(1) All original items whose stems were allophonic homonyms (PA, QA) were discarded. This was done because the observed difficulty of such items was excessively high, because there was concern that these items reduced the homogeneity of the instrument, and because subjects may have been misinterpreting the instructions with regard to these items by seeking homophones which matched their several pronunciations.

(2) All items on the original list which exhibited extreme difficulty (0 to 3, or 13 to 16 errors) were discarded. The investigator wished to maximize reliability and validity by maximizing the variance

of test scores; to accomplish this, an attempt was made to have all item difficulties in the neighborhood of 50%. While it is desirable to include some very easy and very difficult items in order to discriminate between subjects with extreme ability scores, the investigator was confident that such items would be present in any event, owing to the fallibility of his procedures for selecting items of 50% difficulty.

(3) The proportion of positively-keyed items was increased from 50% to 75%, since negatively-keyed items had apparently contributed almost nothing to the variability of pilot study scores. (The reasoning behind this decision subsequently appears to have been superficial.)

(4) Using the graphic regression of TL rank upon difficulty for Pa and Qa items, as exhibited in Figure 4.1 above, 22 Pa and 14 Qa items were selected. The distribution of each set of TL ranks was centered near the 50% difficulty point on the appropriate regression line.

(5) Only hNa words were used as negatively-keyed items. This was in the hope of replicating the pilot study success of SPOKE.

The selection of both positively- and negatively-keyed items was again non-random, and performed by the investigator. In this way it was hoped that the validity of the regression procedures would be maximized, since any variables correlated with TL rank or with accidental parallelism in his selection procedures would continue to operate, even if the true relationship between the explicit selection variables and item difficulty were quite tenuous. It was assumed that the new group of subjects would be of the same ability level as the old, and that item difficulty was independent of the proportion of positively keyed items.

These procedures resulted in a list of 48 items, of which 36 were homophones and 12 were not. Only five items were retained from the previous list (MIGHT, PEER, QUARTZ, BOULDER, SPOKE). The hNa word PRESENT, from the verse selection in the pilot study, was included, as was ALOUD, whose homophone ALLOWED had appeared on the same verse instrument. Four minutes was set as the time allowance for the Homophone test.

The second new instrument, Bliss, was based upon the earlier Alexander Pope selection, and required that words having homophones be underlined in a passage of verse. The test consisted of five iambic pentameter couplets, and contained 78 homophones (out of 88 different words). This very high density of homophones was prompted by the apparent inefficiency of negatively keyed items in the pilot study. Since the resulting selection bore very little resemblance to Pope, or anyone else, it was attributed in the instructions to one P. P. Bliss, a historical figure whose death is commemorated in one of the poems of Julia Moore. This deception was motivated by a desire to preserve the illusion of semantic content. All of the homophones from the earlier version which had difficulties between 25% and 75% were included, as well as some previously untried homophones. Four minutes was allotted for this task. While it may seem strange to assign equal time limits to an 88- and to a 48-item test, pilot study subjects had reached line 16 (median) of the earlier verse task by four minutes. The Homophone test would be administered prior to Bliss, so that subjects would have already formulated strategies for identifying homophones; and it was anticipated that the greater difficulty of items in context would leave subjects ample time to record their positive responses.

The third new test, Homonym, was an alternative revision of the list of words used in the pilot study. The task was to underline words which had more than one meaning from among a list of 48 isolated words. Fifteen items from the earlier list were retained. The remaining 33 items were selected in such a way that the test included 37 positively- and 11 negatively-keyed items. Of the latter, six were of the form hna and five Pna, Qna being neglected on theoretical grounds. (It was hypothesized that Hna distractors would be more effective than hna, since some subjects would apply homophone strategies to Homonym items, but that Qna homophones would be perceived as hna because of their greater obsered difficulty as homophone items.) Positively-keyed items were of the form PNa, QNa, hNa, PNA, QNA, and hNA. Table 5.2 lists the frequencies of each type of item.

Table 5.2

FACTORIAL STRUCTURE OF HOMONYM TEST

Homophone Status:		Positively-Keyed (N)				Negatively-Keyed (n)			
		H		h		H		h	
		P	Q		(total)	P	Q		(total)
Allophone Status :	A	2	5	8	15	-	-	-	0
	a	3	1	18	22	5	0	6	11
Totals:		5	6	26	37	5	0	6	11

An insufficient number of possible items having the characteristics A and n exist for a systematic factorial design to be practical here. Ambiguity regarding the expression "more than one meaning" was supposedly avoided by selecting N-type items which function as different parts of speech, or homographs with different etymologies, and contrasting these with n-type items which have only one definition in WCD. No scheme for predicting item difficulty on the basis of the pilot study was discovered, so that the investigator selected items for the Homonym test in a comparatively unsystematic fashion.

No test of allophones was included in this battery, since their inclusion in the pilot study had been sufficient to establish that G → A in the models was generally taken to be unique, even though words having allophones could be recognized when the same subjects were instructed to do so. The small variability in the pilot study data offered relatively dim prospects for quickly revising the original list into a reliable test of allophone recognition.

In addition to the ten reference tests and three experimental tests, a one-page status questionnaire was included in the battery. This requested the name, birthdate, university status, sex, first language, second languages, region where English was first learned, geographic history, and course section of each respondent. This questionnaire, together with the Homophone, Bliss, and Homonym tests, is exhibited below in Appendix C.

Subjects

With the cooperation of the instructors, the investigator solicited students enrolled in three large sections of a summer school course in Educational Psychology at the University of Wisconsin. The study was advertised as a linguistic experiment, involving no electric shocks. Participation was encouraged by a course requirement that all students serve for three hours as experimental subjects. Assistance in the present study satisfied half of that requirement. Eighty-nine non-random volunteers completed the test battery. Since the investigator wished that this group be as homogeneous as possible with respect to language affiliation, only the scores of native speakers of Mid-western American English who were not fluent in another language were included in the analysis. Scores of individuals who did not meet this requirement (which had been stated clearly in the call for subjects), or who had lived for more than a few months outside of the Midwest, according to the questionnaire data, were not analyzed, though these persons were given credit for experimental participation. This exclusion reduced the effective size of the group to 70 persons (28 male, 42 female). The median age was 24.0.

Despite the absence of random sampling, the investigator believes that this "grab group" was representative of secondary-school teachers who are native speakers of Midwestern American English, and who had exposure to Educational Psychology courses in the late 1960's. The last qualification is not facetious; several of the subjects, in conversation with the investigator, indicated some knowledge of Guilford's "Structure of Intellect" theory.

Administration

Each set of tests and instructions was bound in two booklets. The first contained, in order, Plot Titles, Verbal Comprehension, Homophone, Bliss, Omelet, Disemvowelled Words, Homonym, Suffixes, and the status questionnaire. The last seven instruments were reproduced by a photo-offset process from locally typed copy, with the three experimental tests in primer typeface. Variously colored stock was used to separate instruments, in order to reduce the likelihood of an individual using the wrong page of a booklet, and to improve the appearance of the battery. The second booklet contained, in order, Ideational Fluency, Word Beginnings, Advanced Vocabulary, Consequences, and Word Fluency. Of these, only the Advanced Vocabulary was locally typed. Subjects were instructed to record all responses directly on the pages containing the items, rather than upon separate answer sheets; verbal instructions to this effect, and to attempt every item, overrode the printed instructions in the case of Verbal Comprehension. While this non-use of machine-scoreable answer sheets greatly increased the labor of data collection, it minimized the component of psychomotor abilities that might have been important in the scores of the older subjects, and avoided entirely the prospect of invalid scores being collected through individual failures to correlate properly items with answer sheets.

The battery was administered on two occasions to disjunct subsets of the subject group, between 11:30 a.m. and 1:00 p.m. Since one of the sections met in the same classroom until 11:25, participation involved a minimum of initiative, but, in many cases, a late lunch.

Subjects attended whichever session they wished; on Wednesday 32 appeared, and on Thursday 57. Four of the first group, and fifteen of the second were excluded for the linguistic reasons stated above. The investigator, with one assistant, administered the tests to each sub-group and enforced the time limits. Because both of the vocabulary tests (Verbal Comprehension, Advanced Vocabulary) appeared to be scaled well below the ability level of the group, three and six minutes, rather than the recommended four and eight, were alloted to these tests, in deference to the subjects' evident desire to escape from the testing situation. Very few subjects appeared to be still engaged in solving these test items when instructions were given to proceed to the next instrument. Only two, rather than three, Ideational Fluency tasks were given, to avoid exceeding the announced duration of testing. A few subjects who arrived late were administered the Plot Titles test after the other subjects had completed the battery. The only intermission followed the Suffixes test, during which the second booklet was distributed.

Scoring

Item responses for all tests except those marking divergent production factors were recorded onto computer coding sheets, in a format acceptable to the FORTAP item analysis program (Baker and Martin, 1968). The remaining tests were scored by the investigator, using the respective scoring guides, with the exception of Consequences, which was scored by an individual who had had prior experience in scoring this instrument. Part scores, for the purpose of obtaining

Spearman-Brown reliability estimates, were recorded for each part of *Word Beginnings*, *Word Fluency*, *Ideational Fluency*, *Consequences*, and *Plot Titles* tests. The last two instruments yielded two scores each, distinguishing "obvious" and "remote" consequences in the first case, and "low quality" and "high quality" (or "clever" and "non-clever") titles in the second. It was not intended that the "remote" and "high quality" variables be included in the analysis, since there was no reason to expect the experimental tests to be loaded on the DMT (originality) factor which these variables mark, but the scores were recorded so that their correlations with the other measures could be inspected. The FORTAP routine was used to obtain total scores on the *Omelet*, *Disemvowelled Words*, *Verbal Comprehension*, and *Advanced Vocabulary* tests; separate scores were obtained for each half of the last. A typographical error occurred in reproducing one option of an *Advanced Vocabulary* item, which was called to the attention of the subjects at each testing session. The item was retained after item analysis failed to reveal any discrepant features in the item characteristics of that option. Scoring of the experimental tests (*Bliss*, *Homophone*, *Homonym*) is described in the next section. Positive and negative item responses were identified by the use of manual scoring stencils, and coded for analysis by the FORTAP routine.

Preliminary Results

Scores and subscores for all tests were punched, verified, and submitted as input for a standard numerical analysis routine, DSTAT1 (Wetterstrand, 1969), using the CDC-3600 computer. Table 5.3

Table 5.3

MEANS AND STANDARD DEVIATIONS BY SEX, AND RELIABILITIES OF VARIABLES

Variable	Males (N=28)		Females (N=42)		Total (N=70)		Reliability[a]
	Mean	s.d.	Mean	s.d.	Mean	s.d.	
1. ASF-DSU	18.79	6.06	21.12	5.73	20.19	5.94	–[b]
2. BPR-DSU	27.50	-.82	27.64	6.92	27.59	6.83	.70
3. CFL-DSU	50.36	10.78	66.62	10.49	54.11	10.97	.73
4. AIF-DMU	23.00	6.73	29.64	8.65	26.99	8.54	.74
5. BOC-DMU	8.93	3.59	10.62	4.06	9.94	3.94	.64
6. RMCSQ	11.00	4.53	12.00	4.76	11.60	4.66	.55
7. CPT-DMU	10.04	4.96	11.81	6.43	11.10	5.91	.83
8. AGZ-CMU	18.04	3.72	17.93	3.26	17.97	3.43	.77
9. BVI-CMU	11.86	3.12	11.05	2.58	11.37	2.81	.73
10. CVJ-CMU	11.89	2.67	12.05	3.31	11.99	3.05	.71
11. AOM-CSU	17.86	5.54	18.76	6.11	18.40	5.86	.87
12. BDV-CSU	12.57	4.53	14.05	4.61	13.46	4.61	.81
13. BLISS	38.86	15.75	43.62	13.24	41.71	14.38	.95
14. HPHON	20.64	8.27	23.14	6.19	22.14	7.14	.87
15. NOPHO	1.79	1.71	1.48	1.49	1.60	1.57	.53
16. HMNYM	25.68	5.81	27.10	4.47	26.53	5.06	.77
17. NONYM	1.64	1.81	1.17	1.25	1.36	1.50	.55
- BLISS[c]	43.14	16.29	48.55	13.68	46.39	14.90	.94

[a] Values are Hoyt reliabilities, except for variables 2 through 7 where Spearman-Brown formula was used to estimate reliability.
[b] No estimate available.
[c] Includes nine items which were discarded after item analysis.

shows the means and standard deviations for each instrument, subdivided by sex and for the total sample. Using the FORTAP routine on the CDC-1604 computer, item analyses were performed on all tests having more than three items, yielding Hoyt reliabilities and certain item statistics. Reliabilities of the divergent production tests (with the exception of Suffixes)were estimated by the formula

$$R = \frac{kr}{1 + (k-1)r}, \qquad (5.1)$$

for the r the mean correlation between parts, and k the number of parts. For k = 2 this is equivalent to the Spearman-Brown formula. For k = 3, it is more difficult to justify, since the mean of three sample product-moment correlation coefficients is not an unbiased estimate of the population value. The formula yields reasonable results, however, and has the advantage of computational ease; see Houston and Otto (1968), footnote. Although no reliability estimate was available for Suffixes, its correlations with other tests (e.g., r = .613 with Bliss) imply that it had reasonably high reliability. Not shown in Table 5.3 is the "high quality" Plot Titles variable. Its reliability, estimated by (5.1), was only .35, and its correlation with "remote" Consequences, the other DMT marker, was .197, a value not different from zero at the 90% level of significance. Its highest correlation with any other variate was .356, with Disemvowelled Words (a DSU marker). On the bases of low reliability and uninterpretability, this variate was excluded from subsequent analyses.

The Verbal Comprehension test had a mean of 17.97 out of 24 items; a more difficult test would have been appropriate for this group. This caused an effective reduction in its length, since the first nine items showed essentially no variability, contributing to the fairly low (Hoyt $r = .77$) observed reliability. Although the *Advanced Vocabulary* test was also somewhat improperly scaled (a mean of 23.36 out of 36 items), the latter test was entered as two variables (9 BVI-CMU, 10 CVJ-CMU) in the analyses, corresponding to the separately timed halves of this test. This resulted in three CMU markers of similar reliabilities (.77, .73, .71), rather than two disproportionate tests (.77, .91). Because all three were of identical format, less specific variance was introduced into the anticipated CMU factor than would generally result from the splitting of a reference test.

The other reference tests with ceilings, *Omelet* and *Disemvowelled Words*, were well scaled for the group, with mean item difficulties of 61% and 54%, respectively. Except for the first two CMU markers, women's scores had higher means than men's; this may have been a selection artifact. Using a *t*-test with 68 degrees of freedom and assuming homogeneity of variances, the investigator found that the difference between sexes on *Word Fluency* ($t = 2.41$, $p > .95$) and *Ideational Fluency* ($t = 3.38$, $p > .99$) achieved statistical significance.

When the total number of homophones underlined was used as the score for the *Bliss* test, the mean score was 59%, an unexpectedly high value, but satisfactorily close to 50%. Reliability was .94 (Hoyt), easily the highest among all the instruments in the battery. When nine items having poor item characteristics (point-biserial correlation with the total test score between +0.08 and +0.27) were omitted from the scoring, reliability rose slightly to .95.

The efforts to improve the homophone list appeared to have been quite successful. When the number of Homophone items passed was used as the score, the mean score was 68% of the 48 items possible, and the Hoyt reliability was .86. Because the negatively-keyed items appeared to correlate poorly with total test scores (the median point-biserial correlation being .41 for positive, and .18 for negative items), various scoring schemes were investigated. By splitting the score into positive and negative item totals, the reliability of the positive part rose to .87 (despite the loss of 25% of the items). Scored separately, the eleven negative items had a median point-biserial correlation of .36, and a Hoyt reliability of .53.

An alternative scoring scheme was investigated, based upon differential item weights. Using 31 positive and two negative items, and assigning weights ranging from one to four to correct responses, a scale with a median point-biserial correlation of .47, and the reliability of .89 was devised. Item weights were impressionistically based upon the magnitude of the original point-biserial correlations.

A third scoring scheme, based upon a "correction for guessing" formula was also explored. According to the rationale of this sytem, subjects who gave false positive responses were using some fallible strategy such as (3.12) on some of the items; such a strategy would tend to enhance their scores, since most of the items were positively-keyed. By multiplying the number of correct positive responses by the number of correct negative responses, scores were reduced in proportion to the number of false positive responses. No reliability estimates of these scores were made, but scores computed by all three methods were intercorrelated with the other measures.

The Homonym test was less successful than the Homophone test, perhaps because no procedure for item selection was available having the validity of that which was used for the latter test. The average score was 75% of the 48 items correct, and the Hoyt reliability was .77 for all items. Separating this total into a positively- and a negatively-keyed half as above resulted in a 37-item positive part with a Hoyt reliability of .77, and an eleven-item negative part, with a reliability of .55.

A second scoring procedure, using differential weights on a subset of the items (33 positive, five negative), by the same procedure as was used on the Homophone test, increased the reliability of .81. In the light of the extensive literature on the futility of differential weighting (see the review by Stanley and Wang, 1968), this was a rather substantial gain. The weighted scores correlated only .83 with the 37 equally weighted positive items.

A third scoring system, using the product of correct positive and negative items, was also applied to the data, under the same rationale as for the Homophone test.

Intercorrelations between scores on the reference tests with scores based upon the total correct among positive items, and with scores based upon the differential weights derived from the item analyses, are compared in Table 5.4 for the Bliss, Homophone, and Homonym tests. For each pair of correlations between a given test and the unweighted and the derived experimental test scores, that with the larger absolute value is underscored.

Inspection of Table 5.4 shows that for the Homonym test, the derived scores were less highly correlated with the DMU and CMU tests than were the raw scores, while the DSU and CSU markers showed mixed results. The derived scores were less highly correlated with the experimental tests. For the Homophone scores, results were mixed for the divergent production tests, but raw scores shared more variance with the cognitive tests. Raw Homophone scores were more highly correlated with Homonym scores. The Bliss raw and derived scores, although almost perfectly correlated (r = .998), showed less equivocal results; except in the case of the Ideational Fluency test, the derived scores showed slightly higher correlations.

Since these results indicated that the item analysis procedures had somewhat lowered the communalities of the Homophone and Homonym instruments, it was suspected that the derived scores had raised the reliabilities by capitalizing on specific and error variance. The derived Bliss scores, on the other hand, appeared to have reduced the attenuation due to unreliability of the observed correlations. Accordingly, it was decided to use the raw scores (i.e., total positive items passed) for Homophone and Homonym, and the derived Bliss scores, in the subsequent factor and component analyses. This choice of variates permits the inclusion of negative item scores from the first two instruments (called NOPHO and NONYM) without introducing problems of item overlap.

No justification is offered for using raw scores rather than products of positive and negative scores, the third scoring methods

TABLE 5.4

Correlations of Raw and Derived Homophone and Homonym Scores
(Decimals omitted; N=70)

c*	Variable	BLISS	$BLISS^d$	HPHON	$HPHON^d$	HMNYM	$HMNYM^d$
(1)	ASF-DSU	607	613	567	563	350	417
(2)	BPR-DSU	492	507	449	460	555	448
(3)	CFL-DSU	487	491	454	455	401	414
(4)	AIF-DMU	162	157	203	231	341	312
(5)	BOC-DMU	075	079	-115	-120	-052	007
(7)	CPT-DMU	-254	-270	-242	-232	-111	-009
(8)	AGZ-CMU	285	307	465	429	491	418
(9)	BVI-CMU	273	287	387	338	363	295
(10)	CVJ-CMU	352	365	430	391	396	323
(11)	AOM-CSU	683	703	618	586	496	563
(12)	BDV-CSU	708	722	663	626	565	553
--	BLISS	945^a	998^b	777	779	604	590
(13)	$BLISS^d$	998^b	948^a	784	783	607	594
(14)	HPHON	777	784	873^a	973^b	590	544
--	$HPHON^d$	779	783	973^b	887^a	583	518
(16)	HMNYM	604	607	590	583	772^a	831^b
--	$HMNYM^d$	590	594	544	518	831^b	806^a
(15)	NOPHO	-179	-186	-226	-242^b	-204	-131
(17)	NONYM	-225	-242	-312	-316	-164	-281^b

[a] Value is Hoyt reliability.
[b] Correlation of measures with overlapping items.
c*This column shows code number of variate in factor analyses.
[d] Indicates score derived from item analysis.

for the Homophone and Homonym instruments. Correlations were respectively .92 and .81 between these alternative measures, and the decision to use the former was arbitrary. The products showed consistently higher correlations with the CMU reference tests, and generally higher correlations with the symbolic tests, compared to the raw scroes. In Chapter VI below are shown the results of a factor and a component analysis of the present data using product scores as dependent variables (HPHOCG, HNYMCG) in place of HPHON, NOPHO, HMNYM, and NONYM, the separate positive and negative scores.

Intercorrelations of all variables, except for the three experimental scoring methods discarded on the basis of the data in Table 5.4, are exhibited in Table 5.5.

Of the factor reference tests, the three CMU instruments held together best, with intercorrelations of .74, .77, and .78. The two CSU tests correlated .70, a reasonably high value. The three DSU measures showed more dispersion, with intercorrelations of .44, .53, and .61. The highest of these values was for Word Fluency and Suffixes, the lowest for Word Fluency and Word Beginnings. The last test correlated .61 with the product score on the Homonym test, and in the .44 to .55 range with the CSU, CMU, and experimental tests. The three DMU tests appeared to be in very poor agreement, with intercorrelations of .18, .15, and .16. For N = 70, an observed correlation coefficient must have a value greater than .201 for the probability of its true value being greater than zero to exceed 90%. The highest correlation between any of these tests and another measure

TABLE 5.5

Intercorrelations Among Twenty-One Variables in Homophone Study (N = 70)*

Variable Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 ASF-DSU																				
2 BPR-DSU	53																			
3 CFL-DSU	61	44																		
4 AIF-DMU	27	33	49																	
5 BOC-DMU	10	07	26	18																
6 RMCSQ	22	34	32	53	03															
7 CPT-DMU	-12	-07	-04	15	16	13														
8 AGZ-CMU	40	55	20	15	-14	29	-22													
9 BVI-CMU	35	44	-00	05	-24	28	-18	77												
10 CVJ-CMU	30	45	00	05	-14	26	-20	78	74											
11 AOM-CSU	52	48	48	13	07	13	-36	41	30	36										
12 BDV-CSU	50	55	42	19	15	13	-15	38	34	41	70									
13 BLISS	61	51	49	16	08	10	-27	31	29	37	70	72								
14 HPHON	57	45	45	20	-11	10	-24	47	39	43	62	66	78							
15 NOPHO	-17	-27	-27	-22	-18	-19	06	-42	-33	-31	-36	-34	-19	-23						
16 HMNYM	35	55	40	34	-05	12	-11	49	36	40	50	57	61	59	-20					
17 NONYM	-40	-33	-40	-12	-03	-16	09	-38	-19	-30	-39	-23	-24	-31	23	-16				
18 HPHOCG	54	47	49	26	-03	14	-24	53	44	47	67	70	71	92	-57	56	-35			
19 HNYMCG	50	61	52	29	-03	17	-15	58	38	45	57	55	59	61	-29	81	-70	61		
20 AGE	01	02	33	06	04	15	-16	02	-18	-15	35	19	19	22	-29	-01	-33	28	20	
21 SEX	-19	-01	-28	-38	-21	-11	-15	01	14	-03	-08	-16	-16	-17	10	-14	16	-17	-17	10

* Decimals omitted

was .53, between Ideational Fluency and Remote Consequences, a finding of obscure significance. The former DMU marker correlated .49 with Word Fluency (DSU), with which "obvious" Consequences had its largest correlation, .26. The third DMU marker, "non-clever" Plot Titles, had as its largest correlation a value of .36 (negative), with the Omelet (CSU) test.

The experimental tests showed fairly high correlations with several of the marker tests. Bliss correlated .61, .51, and .49 with the DSU tests, or somewhat higher than these tests correlated among themselves. With the DMU markers, Bliss correlated .61, .08, and -.27; with CMU, .31, .29, and .37; and with CSU .70 and .72. The small correlations with the CMU markers were encouraging, since the investigator had been concerned that vocabulary size would be a major determinant of homophone recognition, rather than the more interesting fluency factors; but the moderately large correlation with both the DSU and CSU markers suggest that these factors were not well differentiated in the present data.

Variable HPHON, positive Homophone items, showed a similar pattern, with lower loadings on DSU and CSU, and rather higher loadings on CMU. Negative Homophone Items missed, NOPHO, showed small negative correlations with the divergent production markers, but moderate negative values for CMU (-.42, -.33, -.31) and CSU (-.36, -.34); recall that these values are quite attenuated by the low (.52) reliability of this scale. The product score for the Homophone test (HPHOCG) correlated .54, .47, .49 with the DSU markers, .53, .44, .47 with CMU, and .67, .70 with the CSU tests; these values are somewhat higher than for HPHON.

Variable HMNYM, positive Homonym items, had somewhat lower correlations with the DSU markers (.35, .55, .40), and near zero correlations with the DMU markers, with the exception of Ideational Fluency (r = .34). This was the largest correlation of any experimental test with a test associated with DMU, despite the expectations of the investigator that recognizing homonyms would require the divergent meanings be produced. Correlations with the CMU markers (.49, .36, .40) were about the same as those of HPHON. Correlations with the CSU markers (.50, .57) were lower than for the Bliss and Homophone scales. NONYM, the number of false positives, had larger negative loadings on the CSU markers (-.40, -.33, -.40) than did NOPHO, and slightly smaller loadings than NOPHO on the cognition scales. The "corrected for guessing" product scale HNYMCG calculated as (HMNYM x (11 - NONYM)), showed conspicuously higher correlations with the DSU (.50, .61, .52) and CMU (.58, .38, .45) reference tests than did HMNYM, and slightly stronger association with the CSU tests (.57, .55).

Comparatively large intercorrelations were found among the experimental tests. As might be anticipated, Bliss and Homophone (HPHON, NOPHO, HPHOCG) were more similar (r = .78, -.19, .71) than Bliss and Homonym (for HMNYM, NONYM, HNYMCG , r = .61, -.24, .59, respectively); other correlations of interest were .59 (for HPHON and HMNYM), .61 (for the two product scales), and .23 (for NOPHO and NONYM).

Two status variables, AGE (= year of birth) and SEX (0 = female, 1 = male) were included in the correlational analysis. Pointbiseral correlation coefficients with an absolute value greater than .25 occurred between SEX and Word Fluency (r = -.28), Ideational Fluency (r = -.38), and the differently weighted Homonym scores (r = -.26).

Correlational coefficients greater than .25 were observed between AGE and Word Fluency (.33), Omelet (.35), NOPHO (-.29), and NONYM (-.33). That is, younger subjects were superior on each of these measures, the last two being error scores.

When item data from the Bliss, Homophone, and Homonym tests were compared with the pilot study data, it was found that items which were repeated from the pilot study tended to show lower error rates for all three experimental tests. Of the first 24 homophones occurring on both Bliss and on the earlier verse selection, an average of 15.32 were passed on Bliss, versus 11.64 on the pilot study. Of the five overlapping Homophone items, a mean of 3.54 were passed on the present data, versus 2.45 for the pilot study list. Of the ten overlapping homonyms, 7.49 were passed on the Homonym test, against 5.19 in the pilot study.

Three possible explanations of the reduced difficulties of these items seem plausible. The subjects in the main study may have been more capable. Alternatively, statistical regression may have operated, since overlapping items were selected on the basis of their observed difficulty in the pilot study data. If these were items of relatively high observed difficulty from an unspecified population of easy items, the observed results could be expected. A third, and perhaps the most convincing explanation is that dependencies exist among a set of items on the same test, so that a high density of positively-keyed items of moderate difficulty increases the likelihood of a positive response. Audial and peripheral visual clues should inform subjects in a group-testing situation that many positive responses (underlinings) were being made on these tests. Subjects with no intention of cheating are likely to be influenced by this information, and to persist longer in their search for homophones and homonyms, and to respond positively in cases of uncertainty. Supporting this view is the fact that five negatively-keyed items on the Homophone test were taken from the pilot study list. If a similar statistical regression mechanism operated here, or if the ability level were higher for the present group, subjects would tend to fail fewer of these items; whereas if only a different response set (i.e., to underline freely) were present, subjects would fail more of these items. They failed an average of 0.53, versus 0.18 in the pilot study.

Summary of Preliminary Findings

Item analysis and correlational analysis showed that the experimental tests were of relatively high reliability, compared to other verbal tests requiring comparable testing time. The experimental tests showed moderately high correlations among themselves and with the instruments designated as reference tests for Guilford's factors DSU and CSU, with smaller correlations with CMU. Guilford's CMU and CSU factors appeared to be well represented, with DSU markers showing more dispersion and some tendency to correlate with the CSU measures. Factors DMU and DMT were not apparent in the correlation matrix.

Chapter VI

COMPONENT AND FACTOR ANALYSIS OF THE DATA

Since Spearman's (1904) original formulation of intelligence as the first centroid factor of his battery, both the number of factors postulated by mental test theorists and the variety of mathematical techniques proposed for their extraction have proliferated. While the relative merits of dividing the cognitive domain into nine "Primary Mental Abilities" (Thurstone, 1938), of dicing it into 120 structural elements (Guilford, 1967), or of being content with Spearman's primordial _g_ would appear to be a straightforward empirical question, the psychometrician is in fact faced with an embarrassment of methodological options: component analysis, image analysis, and several techniques of factor analysis, each in combination with various rotational alternatives. Any choice among these can be defended, but each yields results which are different, to some degree, from what another method would provide; and these differences are not necessarily small.

Determination of Comparable Common Factors

The present attempt to specify the factors involved in homophone recognition will deal with the diversity of analytical sects in the "ecumenical" manner proposed by Harris (1967). He recommends that the common factors of a set of data be determined by using several computing algorithms to obtain initial solutions, subjecting each solution to orthogonal and oblique rotations, and comparing the different results. Only factors which are "robust" across methods are taken as important substantive results.

To use this strategy, three factoring procedures were employed. The first, an Incomplete Principal Components Analysis (Hotelling, 1933), factors the observed correlation matrix R into eigenvectors (Q) and eigenvalues (B):

$$R = QBQ' \tag{6.1}$$

The analysis is incomplete in that only the m eigenvalues greater than 1.00 ($\overline{B}$) and their associated eigenvectors ($\overline{Q}$) are retained, yielding the number of components specified by Guttman's weak lower bound (1954). Loadings of the observed variables on these components are given by

$$F = \overline{QB}^{\frac{1}{2}} \tag{6.2}$$

This approach is distinct from factor analysis, in that the components attempt to reproduce the total observed variance, rather than distinguishing between common and unique portions.

A second procedure, Harris' $R\text{-}S^2$ Factor Analysis (1962), operates upon a correlation matrix rescaled by estimates of uniqueness. Squared multiple correlations of each observed variable with the others are used as communality estimates, and the number of factors is the number of eigenvalues of the rescaled correlation matrix greater than one.

A third method, Unrestricted Maximum Likelihood Factor Analysis (Jöreskog, 1967), was also applied to the data, providing a factor solution having a statistical, rather than psychometric basis. Here factor loadings and communalities are estimated by Lawley's method of maximum likelihood, and the number of factors determined by a chi-square test of goodness of fit (Lawley and Maxwell, 1963).

Calculations of the Incomplete Principal Components and Harris R-S^2 Factor Analyses were performed on the CDC-3600 computer, using the FACTORI program (Wetterstrand, 1969). The same machine was used for all of the analyses described in this chapter. The Unrestricted Maximum Likelihood Factor Analysis was performed by program UMLFA (Jöreskog, 1967). Initial solutions obtained by these three methods were used to obtain both orthogonal and oblique derived solutions, based upon Kaiser's (1958) normal varimax procedure, and Harris and Kaiser's (1964) independent cluster solution. The varimax rotation was performed by the FACTORI program; the oblique by program OBLIQUE (Wetterstrand, 1968).

In discussing the results of these analyses, Guilford's practice of regarding as significant only factor loadings greater than or equal to .30 will be followed. The term "common factor" denotes any factor or component having more than one significant loading.

The procedures listed above were applied to two subsets of the intercorrelations appearing in Table 5.5. The first of these was the intercorrelations of variables one through 17, hereafter Matrix A. The second of these consisted of the intercorrelations of variables one through 13, and 18 and 19, the difference being that the "corrected for guessing" product scores were substituted for the separate positive and negative item scores of the *Homophone* and *Homonym* tests.

For the 17 variable Matrix A, the Incomplete Principal Components (IPC) and Harris' R-S^2 orthogonal solutions each obtained five common factors, and the Unrestricted Maximum Likelihood (hereafter UMLFA) orthogonal solution obtained three. The fifth common component corresponded to the sixth largest eigenvalue, and six components were rotated, yielding six common oblique components. Oblique rotation did not change the number of common Harris R-S^2 or of UMLFA factors.

For the 15 variable Matrix B, all methods yielded three orthogonal and oblique factors, which resembled the first three common factors obtained from Matrix A.

There had been some concern that the small number of observations and the skewness of the distributions of the NOPHO and NONYM variables would adversely affect the validity of the chi-square test of a number of factors in the UMLFA analysis. Using the Matrix A data, for the number of factors *k* = 2,3,4,5, and 6, the UMLFA program associated significance levels of .00, .24, .54, .68, and .94, respectively; these values are interpretable as estimates of the probability that the true number of factors is less than *k* (Joreskog 1967, p.458). For Matrix B, *k* = 2,3,4, and 5 were associated with probabilities of .00, .57, .90, and .93. To guard against underfactoring, but at considerable risk of overfactoring, solutions for *k* = 3,4, and 5 for Matrix A, and *k* = 3 and 4 for Matrix B, were retained for use in subsequent comparisons. (Generally the minimal number of factors, three in each of the above cases, for which the probability exceeds 0.05 is retained.)

The next step in the determination of robust common factors is the comparison of the different solutions. A variable is "relevant" to a factor in a given solution if it has a significant loading on that factor. A "comparable common factor" (CCF) is a factor having at least two of the same relevant variables for at least four of the six solutions, with the unions of the orthogonal and of the oblique UMLFA factorizations each regarded as single solutions for these purposes. A variable is "relevant" to a comparable common factor if it is relevant to at least five of the common factors defining the CCF. A variable is "non-relevant" to a CCF if it is relevant to fewer than four of the defining common factors. By these rules, a factor relevant to a CCF must appear in both orthogonal and oblique derived solutions.

Two other types of factors are of interest in comparing solutions. A "comparable specific factor" (CSF) has only a single relevant variable common to at least five of the six solutions. A "non-comparable factor" is a factor having no variable in common with at least five of the six solutions.

Matrix A yielded five comparable common factors, two non-comparable factors, and no comparable specific factors. Matrix B yielded three CCF's and one non-comparable factor. The CCF's from Matrix B closely correspond to the first three CCF's from Matrix A, when CCF's are ordered according to the latent roots of the unreduced correlation matrix to which the principal component involved in each CCF corresponds.

Table 6.1 shows the results for Matrices A and B. Variables shown in parentheses are semi-relevant, or define non-comparable factors. Loadings below .30 are not listed. For Matrix A, the four-factor UMLFA solution is not shown for CCF-1 and CCF-2.

No variables had relevant loadings on the factors that corresponded to CCF-1 and CCF-2 in this solution which were not also relevant in either the three- or five-factor UMLFA contributions. Since 15 NOPHO and 17 NONYM are error scores, negative loadings on these variates do not indicate conceptually bipolar factors.

In general, the factors obtained from Matrices A and B are highly robust, with two of the non-comparable factors being rotations of the same initial solution, and the third a result of overfactorization which vaguely resembles CCF-4.

The intercorrelations of oblique factors for the CCF's from both data matrices are shown in Table 6.2. These are all in fairly close agreement, with the exception of the four-factor UMLFA solution for Matrix B, whose factors show negative correlations. This was the least interpretable of all solutions, since it split CCF-2 into two parts, one the only non-comparable factor for Matrix B, the other an atrocious bipolar factor having middling positive loadings on the Advanced Vocabulary halves, and a larger negative loading on Word Fluency.

Description and Interpretation of Factors

In the discussion below, the following expressions will be used to describe the magnitude of factor (component) loadings: near zero (range .00-.29); small (.30-.49); moderate (.50-.65); strong (.65-.79); very strong (.80-1.08). The last upper limit is necessitated by the

OBLIQUE routine, which occasionally returns values greater than 1.00.

CCF-1

This was the largest factor, and quite robust, although oblique rotation tended to reduce the number of relevant variables for Matrix A. It was defined by very strong loadings on BLISS, strong loadings on the CSU markers and on the homophone variables (14 and 18), with strong to moderate loadings on the homonym and DSU scores. The last were generally reduced by oblique rotation in Matrix A, but persisted in Matrix B. Marginal negative loadings on relevant variable 7 can probably be dismissed as sampling error. Both the DSU and CSU factors appear to coalesce in the orthogonal versions of this factor, which resembles the fifth CCF reported by Harris (1969) in a reanalysis of some of Guilford's data. The oblique common factors identified with CCF-1 are more closely related to CSU, but substantial loadings on DSU markers remain in the Matrix B factors. The factor appears to represent fluency in producing words which satisfy formal requirements. Although "fluency" factors are generally associated with tasks involving multiple responses to a given stimulus, these data suggest that the CSU and experimental test items are solved by comparing the stimuli with elements of internally generated lists.

CCF-2

This CCF is the most interpretable of all, being defined by common factors with strong loadings on the three "vocabulary" tests with small or near zero loadings on other variables. The marginal relevance of variable 15 in Matrix A, and of variable 5 in Matrix B, can probably be ignored (particularly in the latter case, where the negative loadings are nonsensical). It closely corresponds to Guilford's CMU factor, although some caution should be exercised in this identification: the three reference tests are identical in format, and some variance specific to this format is surely involved in this factor. It involves the ability to recognize synonyms.

CCF-3

This factor was quite robust across solutions, and was defined by strong loadings on Ideational Fluency (median .75) and Remote Consequences (.66), and by moderate to small loadings on Plot Titles and DSU variables 2 and 3 for both matrices. The first DSU variable was relevant on a few defining factors for both matrices, while the error scores and Homonym product score showed small loadings on solutions derived from the three- factor UMLFA. These variables appear to define a factor resembling Guilford's DMU, alias "Ideational Fluency" (French, 1963), despite the failure of the third DMU marker to load on this factor. The strong (.70-.80) loadings of some of the UMLFA factors on the Word Fluency test suggests some degree of coalescence with DSU. Peculiarities of this instrument in the present data, which may be related to the UMLFA results, will be described in the next section. CCF-3 seems to involve the production of meaningful words, under simple restrictions.

CCF-4

This factor is difficult to interpret, having as its relevant variables two DSU markers and the homonym error score. The Omelet test had minimal relevant loadings on this factor in four solutions. NONYM has the largest loadings. Recall that on CCF-1, the DSU markers had generally smaller loadings in Matrix A than in Matrix B; in the latter NONYM is subsumed under HNYMCG, and CCF-4 coalesces with CCF-1. A possible interpretation is that CCF-4 is DSU, although the absence of the second DSU marker is puzzling. CCF-1 and CCF-4 correlated .60 in all the oblique solutions, and it is tempting to dismiss CCF-4 as a fragment broken from CCF-1 by overfactoring and correlated error. The DSU markers contributed little to the simple structure of the CCF's, with a different pair relevant to each of CCF's 1,3, and 4 for Matrix A, but all three relevant to CCF-1 for Matrix B. (Compare these with the CSU markers, which all loaded strongly or CCF-1, or the CMU markers, loaded very strongly or CCF-2. Two of the DMU markers loaded strongly and moderately on CCF-3, while a third divided between CCF-5 and CCF-2 its common variance, of which it had little.) It is difficult to give CCF-4 an interpretation which contrasts it with CCF-1.

CCF-5

This factor was defined by strong to moderate loadings on Consequences and the NOPHO scores, with small loadings on the (non-relevant) CSU marker. Table 6.1 shows a specific factor for the Harris R-S^2 orthogonal solution; this was rotated, yielding a common oblique factor which corresponded to CCF-5. It is difficult to attach much significance to this factor, since the raw correlation of its relevant variables was -.18, though it was "robust" in the Harris sense.

It should be understood, of course, that the procedure for determining CCF's offers the analyst protection against spurious factors

TABLE 6.1

Factor Results for Homophone Matrices*

	Orthogonal				Oblique			
	I	II	III		I	II	III	
			a	c			a	c
COMPARABLE COMMON FACTOR 1								
MATRIX A:								
1 ASF-DSU	59	55	56	56		63	49	40
2 BPF-DSU	54	46	45	47	36		32	
11 AOM-CSU	73	68	71	69	57	34	72	32
12 BDV-CSU	81	76	74	76	84	52	77	48
13 BLISS	90	89	95	93	92	90	107	94
14 HPHON	81	79	79	79	76	71	81	80
16 HMNYM	74	62	58	60	85		55	55
(3 CFL-DSU)	54	47	41	45				
(5 BOC-DMU)								-31
(7 CPT-DMU)			-30				-34	
(10 CVJ-CMU)			34	30				
(15 NOPHO)						45		40
(17 NONYM)					49			

	Orthogonal				Oblique			
	I	II	III		I	II	III	
			a	b			a	b
MATRIX B:								
1 ASF-DSU	71	57	59	63	69	52	45	52
2 BPR-DSU	59	47	49	53	47	32		42
3 CFL-DSU	70	52	51	56	71	47	31	33
7 CPT-DMU	-34		-32		-45	-42	-42	
11 AOM-CSU	84	75	80	78	92	86	83	75
12 BDV-CSU	81	79	79	80	85	83	82	83
13 BLISS	87	85	88	91	96	95	97	95
18 HPHOCG	80	76	76	77	80	75	71	71
19 HNYMCG	70	58	60	62	65	51	43	47
(6 RMCSQ)						-30	-34	
(8 AGZ-CMU)	37				31			
(10 CVJ-CMU)	30		33	34				32

*Decimals have been omitted. Parentheses indicate non-relevant variables. See next page for key to factor solutions.

TABLE 6.1 (continued)

	Orthogonal			Oblique		
	I	II	III a c	I	II	III a c
COMPARABLE COMMON FACTOR 2						
MATRIX A:						
8 AGZ-CMU	85	81	85,88	87	74	79,89
9 BVI-CMU	87	82	81,81	92	91	80,86
10 CVJ-CMU	86	82	81,82	89	88	79,84
15 NOPHO	-45	-32	-32,-38	-46		
(2 BPR-DSU)	40	42	35,40			
(3 CFL-DSU)				-37	-44	-45,-38
(5 BOC-DMU)				-35		-34,-37
(6 CPT-DMU)	30					
(16 HMNYM)			32			

	I	II	III a b	I	II	III a b
MATRIX B:						
5 BOC-DMU	-46			-55	-38	-36
8 AGZ-CMU	81	82	83,87	79	82	84
9 BVI-CMU	87	81	81,79	88	86	86,56
10 CVJ-CMU	83	82	81,80	83	83	84,61
(2 BPR-DSU)	37	42	35,35			
(3 CFL-DSU)				-41	-39	-41,-73
(6 RMCSQ)	30			30		
(18 HPHOCG)	30	32				
(19 HNYMCG)	33	34	31,31			

Key to factor solutions:

I	Incomplete Principal Component
II	Harris R-S^2
IIIa	UMLFA, three factors
IIIb	UMLFA, four factors
IIIc	UMLFA, five factors

TABLE 6.1 (continued)

	Orthogonal			Oblique		
	I	II	III a b c	I	II	III a b c
COMPARABLE COMMON FACTOR 3						
MATRIX A:						
2 BPR-DSU	34	30	43,31			37
3 CFL-DSU	42	44	80,41,41			80
4 AIF-DMU	80	70	58,78,75	79	59	64,77,73
6 RMCSQ	74	67	46,61,60	72	65	54,61,59
7 CPT-DMU	52		34	60	49	37,41
(1 ASF-DSU)			47			38
(8 AGZ-CMU)			32			31
(15 NOPHO)			-33			-35
(17 NONYM)			-42			-43
MATRIX B:						
2 BPR-DSU	36	30	42,32	32	38	40
3 CFL-DSU	46	48	70,44	42	62	78
4 AIF-DMU	78	69	66,67	80	77	78,59
6 RMCSQ	76	66	54,77	79	64	64,81
7 CPT-DMU	52			58	36	30,34
(1 ASF-DSU)			41		33	39
(19 HNYMCG)			37			33
COMPARABLE COMMON FACTOR 4						
MATRIX A:						
1 ASF-DSU	50	43	44,46	62	44	46,49
3 CFL-DSU	51	46	65,67	67	63	80,81
11 AOM-CSU	34		33	32		33
17 NONYM	-83	-59	-30,-51	-104	-72	-32,-61
(7 CPT-DMU)				34		
(8 AGZ-CMU)			37,30			31
(15 NOPHO)			-53			-56
MATRIX B: (See Noncomparable Factor 8)						

TABLE 6.1 (continued)

	Orthogonal				Oblique			
	I	II	III		I	II	III	
			b	c			b	c
COMPARABLE COMMON FACTOR 5								
MATRIX A:								
5 BOC-DMU	80	57		48	80	64		63
15 NOPHO	-67			-43	-71	-57		-60
(11 AOM-CSU)						39		46
(12 BDV-CSU)				33		44		55
NONCOMPARABLE COMMON FACTOR 6								
MATRIX A:								
(1 ASF-DSU)						-33		
(16 HMNYM)						67		
NONCOMPARABLE COMMON FACTOR 7								
MATRIX A:								
(7 CPT-DMU)		47						
(11 AOM-CSU)		-32						
NONCOMPARABLE COMMON FACTOR 8								
MATRIX B:								
(1 ASF-DSU)							32	
(2 BPR-DSU)							34	
(3 CFL-DSU)			55				48	
(8 AGZ-CMU)			34				85	
(9 BVI-CMU)							35	
(10 CVJ-CMU)							30	
(19 HNYMCG)			33				51	

TABLE 6.2

INTERCORRELATIONS AMONG OBLIQUE FACTORS*

Comparable Common Factor	Matrix A 1	Matrix A 2	Matrix A 3	Matrix A 4	Matrix B 1	Matrix B 2
2-I	46				42	
II	44				50	
IIIa	38				49	
IIIb	41				-10	
IIIc	46					
3-I	18	13			32	12
II	09	10			48	32
IIIa	54	30			56	36
IIIb	20	13			18	-22
IIIc	15	12				
4-I	61	38	29			
II	58	45	46			
IIIb	61	39	42			
IIIc	59	39	37			
5-I	23	08	23	28		
II	65	40	24	58		
IIIc	69	43	22	59		

*Decimals omitted.

Key to solutions:

I	Incomplete Principal Component
II	Harris R-S^2
IIIa	UMLFA, three factors
IIIb	UMLFA, four factors
IIIc	UMLFA, five factors

which are artifacts of a particular factoring or rotational technique, but is no defense against correlated errors in the data. These same two variables (5 and 15) had significant loadings in solutions defining CCF-2, where they had little interpretability.

The interpretation of CCF-4 and CCF-5 is further complicated by the fact that in both cases the highest loading is on a known instrument showing rather poor correlations with the other reference tests for the same Guilford factor; and the second highest loading is on an experimental test of low reliability, having a severely skewed distribution. Whether CCF-4 and CCF-5 would persist if a normalizing transformation were applied to the NONYM and NOPHO variables is an open question.

Table 6.3 shows the uniqueness estimates obtained by the UMLFA and Harris R-S^2 initial solutions, and the variance unaccounted for by the IPC solution for Matrices A and B. These are in fairly close agreement, with the exception of variables 5,15, and 17, and to a lesser extent 4 and 6, which show considerable dispersion. The IPC solution for Matrix A accounted for far more of the variance than did the other solutions; CCF-5 may be an artifact of this idiosyncracy of the IPC analysis, since it has its only relevant loadings on two of these variables, and its only strong loadings on the solutions derived from the IPC analysis.

Sex Differences

Before attempting a more detailed interpretation of these results, the intercorrelations among these variables for men and for women were examined separately, in the hope of finding some explanation of the

TABLE 6.3

Uniqueness Estimates for Matrices A and B*

Initial Solution:	I		II		IIIa		IIIb		IIIc
Matrix:	A	B	A	B	A	B	A	B	A
Variable:									
1 ASF-DSU	343	420	364	423	445	457	438	459	423
2 BPR-DSU	392	387	438	437	481	461	469	468	460
3 CFL-DSU	195	255	324	321	142	214	133	140	118
4 AIF-DMU	286	328	494	547	657	551	341	491	377
5 BOC-DMU	261	655	692	773	866	866	867	879	694
6 RMCSQ	312	333	542	556	741	646	559	357	565
7 CPT-DMU	528	583	732	745	876	834	794	807	773
8 AGZ-CMU	152	182	196	209	120	157	094	028	096
9 BVI-CMU	186	190	293	301	280	266	283	290	227
10 CVJ-CMU	198	212	270	278	237	230	239	233	241
11 AOM-CSU	256	261	328	346	404	331	375	345	302
12 BDV-CSU	225	315	323	327	382	333	381	319	251
13 BLISS	155	223	194	267	082	207	074	175	111
14 HPHON	238		267		295		297		256
15 HOPHO	316		6[illegible]4		767		768		617
18 HPHOCG		281		292		302		303	
16 HMNYM	296		361		516		480		472
17 NONYM	240		641		738		650		648
19 HNYMCG		375		364		405		390	

*Decimals have been omitted.

Key to factor solutions:

I	Incomplete Principal Component
II	Harris $R\text{-}S^2$
IIIa	UMLFA, three factors
IIIb	UMLFA, four factors
IIIc	UMLFA, five factors

For Incomplete Principal Components Analysis (I) figures show proportion of variance not reproduced by components.

dispersion of DSU and DMU markers, and to investigate the wisdom of using heterosexual data. Bereiter (1959) found pronounced differences on fluency factors between bright 10th grade girls and boys. Girls showed a word fluency (DSU) factor loading on the present variables 1 and 3, whereas boys showed a general fluency factor and a specific Suffixes factor. To report correlation coefficients for mixed populations which combine diverse covariance structures is to make misleading or uninterpretable statements. While the separate groups of men (N = 28) and women (N = 42) were too small for separate factor analyses to be advisable, the data presented in Table 6.4 were inspected for gross or systematic differences which might affect the interpretation of the five comparable common factors.

In seven cases no differences (to two decimal-places) were observed, in 53 cases females showed the larger (absolute) correlation, and in 118 cases males showed the larger loading. This suggests that the women had scores which were either more attenuated by unreliability, or that the factors were more widely separated for women. In the latter case, intercorrelations among markers of the same factor should be comparable in magnitude for both sexes. Table 6.5 shows the median correlations by sex within nominal clusters (excluding DMU; as can be seen from Table 6.4, the intercorrelations for both sexes among its reference tests were essentially zero, so there was no reason to include it as a cluster.)

The data in Table 6.5 suggest that for the CMU and CSU markers men's and women's scores were equally reliable but women had more differentiated clusters. For the experimental and DSU tests, the women's scores show lower within-cluster correlations, most conspicuously on the experimental tests.

TABLE 6.4

Intercorrelations of Variables for Female ($N_f = 42$) and Male ($N_m = 28$) Subjects*

Variable	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1 ASF-DSU	52 58	56 62	09 46	05 07	11 35	-30 10	36 47	37 42	23 43	57 44	49 48	60 61	61 50	-16 -15	24 44	-34 -43	61 42	41 56
2 BPR-DSU		34 64	27 51	-04 26	42 21	-13 04	55 55	53 33	40 57	44 46	56 54	53 49	45 46	-29 -25	61 51	-28 -41	50 45	66 60
3 CFL-DSU			39 51	24 17	35 22	-11 -03	08 38	-06 17	-18 31	46 51	33 48	35 62	34 53	-27 -23	25 54	-37 -38	45 49	43 59
4 AIF-DMU				11 11	64 31	07 16	06 37	03 25	-13 47	05 27	05 34	-09 42	-01 40	-33 00	26 42	-01 -15	16 30	19 35
5 BOC-DMU					10 -15	17 05	-31 11	-20 -12	-18 -07	06 06	10 15	15 -10	-11 -23	-22 -07	-17 02	-08 11	02 -20	-09 -06
6 RMCSQ						04 28	26 33	29 33	22 35	13 09	16 04	-00 20	00 19	-28 -05	11 11	-16 -14	13 11	17 15
7 CPT-DMU							-19 -28	-27 01	-27 -07	-44 -25	-22 -10	-47 -03	-47 -01	-07 33	-05 -28	16 08	-37 -14	-14 -25
8 AGZ-CMU								77 77	77 83	37 47	37 41	32 31	53 42	-27 -61	47 53	-34 -44	56 52	57 61
9 BVI-CMU									75 79	30 34	26 52	27 37	36 48	-17 -54	35 43	-24 21	38 57	41 41

*Decimals ommitted.
Upper values are for women.

TABLE 6.4 (continued)

Variable	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
10 CVJ-CMU										30 48	40 46	37 38	37 55	-16 -56	28 59	-17 -48	38 63	32 67
11 AOM-CSU											69 70	71 71	57 70	-33 -39	40 62	-34 -40	65 71	53 63
12 BDV-CSU												71 73	63 70	-24 -45	52 61	-25 -17	66 73	56 53
13 BLISS													70 85	-26 -08	40 79	-23 -22	70 71	47 67
14 HPHON														-12 -31	41 73	-25 -33	89 93	50 68
15 NOPHO															-23 -15	-03 -44	-53 -60	-18 -38
16 HMNYM																02 -28	45 63	77 84
17 NONYM																	-22 -43	-61 -75
18 HPHOCG																		51 68

TABLE 6.5

Median Intercorrelations Among Nominal Clusters
For Female (N_f = 42) and Male (N_m = 28) Subjects*

Variable	Sex	CMU	CSU	EXP	DSU	NEG
CMU	f:	77	33	36	37	-21
	m:	79	47	43	42	-51
	t:	77	37	39	35	-32
CSU	f:		69	60	47	-34
	m:		70	70	50	-40
	t:		70	64	49	-35
EXP	f:			41	45	-23
	m:			78	51	-25
	t:			61	49	-22
DSU	f:				52	-29
	m:				58	-32
	t:				53	-30
NEG	f:					03
	m:					44
	t:					23

*Decimals omitted.

Key to clusters:
DSU = 1,2,3
CMU = 8,9,10
CSU = 11,12
EXP = 13,14,16
NEG = 15,17

f = women
m = men
t = total group
(m + f)

Variable by variable, men's scores show consistently higher correlations for variables 3,4,10,11,14,16,17, and 19. Women's scores show consistently higher correlations for variable 7. Differences of the largest magnitude involved variables 3,4, and 16.

In general, men's scores appear to show more communality. The DSU markers showed less dispersion, and the three experimental tests were much more highly intercorrelated than for women. Variable 4, which was virtually orthogonal to the other variables for women, showed moderate correlations with the CSU markers, and small positive correlations with the CMU and EXP measures for men. The error scores NOPHO showed moderate correlations with all CMU tests for men and near zero for women. It is quite difficult to reinterpret the factor analysis on the basis of these observed differences, or to account for the differences themselves. Had all subjects more resembled the male subgroup of the present study, however, it seems likely that the following changes would have occurred in the results: HMNYM would have had a higher loading on CCF-1, while CPT-DMU would have been non-relevant to this factor; the communality of the first three CCF's would have been somewhat greater; and CCF-4 and CCF-5 would not have emerged as common factors with relevant loadings on the same variables as in the present data.

Factors of the Experimental Tests

The largest loadings of the three experimental tests, apart from the scores on negatively-keyed items, were on CCF-1, a factor strongly related to CSU, and upon which the DSU markers were substantially loaded. Identification of CCF-1 with factors found in previous studies is, of course, an uncertain enterprise. Besides CSU and the fifth CCF of Harris' reanalysis, "Verbal Closure" (Pemberton, 1960) shows some resemblance. The highest loadings on the CCF-1 factor were for the Bliss test scores (homophones in context) although it appeared that for males the Homophone (no context) scores were about as highly loaded. Specific variance in the format of the experimental tests has surely influenced the location of this factor, but the high loadings on the symbolic reference tests suggest that either Bliss or Homophone may be of value to other researchers as reference tests for this factor.

Use of the "corrected for guessing" scores, rather than the raw total, had little effect on the Bliss and Homophone loadings, but shifted the factor so that the CSU and DSU loadings were higher, particularly in the oblique solutions, and for the Omelet test. Specificity due to response set in the Homophone and Homonym scores apparently was reduced, so that this scoring procedure appears to be preferred among the several considered above. Error scores and corrected scores showed small loadings on CCF-2, a CMU or "Verbal Comprehension" factor. This was almost the only vestige of the "Semantic mode operators" hypothesized in Chapter III, although the use of more difficult words would probably have increased the loadings on this factor. The large differences between men and women with respect to false positive error scores is striking: the median correlation for NOPHO with the CMU markers was -.17 for women and -.56 for men. While the small numbers involved make sampling error or differently shaped distributions a plausible explanation, it is also possible that different sexes used different strategies. The Ideational Fluency scores (naming things which fit categories) correlated -.33 with NOPHO for women, and .00 for men; whereas "non-clever" Plot Titles correlated .07 for women and .33 (positive) for men with NOPHO. Though both are DMU markers, the first task involves denotative precision, while the second reflects fluency in expressing aspects of a complex situation, where connotative meaning operates. Fluency in interpreting denotative meaning is relevant to the homophone task according to models (3.10-13), but much less so in (3.14-16). Very generalized fluency, such as that which Bereiter (1959) found in 10th grade boys, in conjunction with a strategy resembling (3.16), could lead to correlations resembling those reported above, particularly if men tended to use (3.16) and women a strategy like (3.10-13). Such attempts at detailed interpretation of course impose considerable strain upon the present corpus of data, and specifically designed experiments could probe such issues more convincingly and at less risk of self-delusion.

The Homonym test was less highly loaded on CCF-1 than the other experimental instruments, being non-relevant in the Harris R-S^2 derived oblique solutions for Matrix A, and generally ranking fifth on this factor, between Homophone and DSU markers. This is not unexpected, since the task minimizes the formal constraints, and emphasizes semantic ones; but it is noteworthy that only small to near-zero loadings appear for this test on the factors defining CCF-2 and CCF-3, which were interpreted as having primarily semantic significance.

This can be rationalized in the following way: CCF-2 includes considerable specific variance related to item format, and reflects vocabulary size. Ignorance (absence in memory) of one or all meanings of the stem or of the keyed option is probably the major cause of failed items, whereas on the Homonym test, failure to recall meanings which could be recognized successfully in a different task probably accounts for most of the variability in the data. The Bliss, Homophone, and Homonym tests all involve a process more like recall than recognition, despite the use of the latter term in deference to the conventional description of the item format. In the tests loaded on CCF-2 the subject matches the stem to another stimulus, whereas in the CCF-1 tests the stem is matched to internally generated (retrieved) information.

CCF-3, on the other hand, resembles "Ideational Fluency," the listing of words semantically related to a stimulus; the symbols are retrieved on the basis of their meanings. The Homonym test involves the opposite task: the symbol is presented as a principle by which denotations are to be retrieved. There is no reason to suppose that the abilities required to perform these two processes are highly correlated, though the UMLFA solutions assign the corrected Homonym scores a small relevant loading on this factor.

Note finally that the variables relevant to CCF-1 all involve a certain degree of semantic restraint upon fluency. Words, rather than strings of letters, are listed in the DSU tasks and are the solutions to the CSU tasks. Thus the results are not difficult to reconcile with the models of Chapter III.

Summary and Implications for Further Research

Most of the common variance of the experimental tests was accounted for by a single comparable common factor. Substantially loaded on this factor were the experimental tests and instruments designated as reference tests for the cognition and divergent production of symbolic units. It resembles most closely a factor obtained by Harris in a reanalysis of some data of Guilford's using similar comparative analytic techniques.

The procedures used for test development appear to have been quite successful, in that both the Bliss and the Homophone tests appear to be useful reference tests for this factor. Bliss has the advantage of an extremely high univocal loading and reliability for subjects like those in the present group; whereas by using the Homophone List of Appendix A and following the procedures detailed in Chapters IV and V, Homophone can be revised and rescaled to accomodate the ability level of a given sample of individuals.

As the absence of a "related research" section in this paper may have suggested, little research by behavioral scientists in the area of homophones has seen publication. Carrol (1941) included a Memory of Homophones test in a factor analytic battery, using college men and women as subjects. This loaded on a factor which he named "Conventional Linguistic Response," and identified as part of Thurstone's W. Other variables relevant to this factor, in increasing order of magnitude, were: (small:) Suffixes, a test resembling Omelet, Rhymes (a divergent production task), and a test of Grammar; (moderate:) Vocabulary (a CMU-type test), and Diction, a test requiring the detection of words having slightly inappropriate meanings for their context. The use of subjective rotation methods in the older studies makes comparison with the present results quite difficult, but it appears that the task was rather different from the present one. Because of their susceptibility to auditory confusion, homophones have been used occasionally as stimuli in memory experiments (e.g., Davis, 1932; Kintsch and Buschke, 1969). The Homophone List in Appendix A, and the hypotheses regarding item difficulties, may be useful in future memory research.

The coalescence of Guilford's DSU and CSU, and the paradoxical factorial simplicity of a task inspiring such complex conceptual models as (3.15), invite further empirical investigations, as do certain sex differences, noted above. The dispersion of the DMU reference tests suggests some defect in their scoring or in the psychometric naivete of the subjects, some of whom may have attempted to appear "creative," or whatever they believed these instruments measured. Some attempt at replication, including perhaps measures of redefinition (NST), might be instructive.

The description of individual differences in terms of a small set of hypothetical variables antedates Hippocrates' theory of the Four Humors; this psychophysiological model maintained some scientific status for two millenia, and its vestiges persist today in words for the affective states, and in the lore of herbalists, astrologers, and alchemists. Even in the search for invariants, however, the present age is one of swifter change, and it seems unlikely that whatever credence science grants to present factor analytic model of intelligence will long endure. While it is difficult to underestimate the importance of homophones and their recognition with respect to the happiness and the survival of man, their study may contribute to an understanding of the fundamental processes of the mind, whose comprehension and facilitation are the common goals of psychology and education.

REFERENCES

Baker, F. B., & Martin, T. J. *FORTAP: A Fortran test analysis package*. Wisconsin Research and Development Center for Cognitive Learning, 1968.

Barnhart, C. L. (Ed.) *The American College Dictionary*. New York: Random House, 1960.

Bereiter, C. *Verbal and ideational fluency in superior tenth-grade students*. Doctoral dissertation, University of Wisconsin, 1959.

Carrol, J. B. A factor analysis of verbal ability. *Psychometrika*, 1941, *6*, 279-307.

Cronbach, L. J. Essentials of psychological testing. (2nd edition) New York: Harper and Brothers, 1960.

Dante, A. (*II convito*) *The banquet of Dante Alighieri, translated by Elizabeth Price Sayer*. London: Roulledge & Sons, 1887.

Davis, F. C. The functional significance of imagery differences. *Journal of Experimental Psychology*, 1932, *15*, 630-661.

Fisher, R. A., & Yates, F. *Statistical tables for biological, agricultural, and medical research*. (4th edition) Edinburgh and London: Oliver and Boyd, 1953.

Fowler, H. W. *A dictionary of modern English usage*. (2nd edition) New York: Oxford University Press, 1965.

French, J. W. The description of aptitude and achievement factors in terms of rotated factors. *Psychometric Monograph No. 5*, 1951

French, J. W., Ekstrom, R. B., & Price, L. A. *Manual for kit of reference tests for cognitive factors*. Princeton, New Jersey: Educational Testing Service, 1963.

Gleser, R. (Ed.) *Training and educational research*. Pittsburgh: University of Pittsburgh Press, 1962.

Guilford, J. P. *The nature of human intelligence*. New York: McGraw-Hill, 1967.

Harris, C. W. Some Rao-Guttman relationships. *Psychometrika*, 1962, *27*, 247-263.

Harris, C. W. On factors and factor scores. *Psychometrika*, 1967, *32*, 363-379.

Harris, C. W., & Kaiser, H. F. Oblique factor analytic solutions by orthogonal transformations. *Psychometrika*, 1964, *29*, 347-362.

Harris, M. *An illustration of a factor analytic interpretation strategy*. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, February 5-8, 1969.

Hoepfner, R., J. P. Guilford, and P. R. Merrifield. A factor analysis of the symbolic evaluation abilities. *Report of the Psychological Laboratory of the University of Southern California,* 1964, No. 33.

Homerus, (*Odyssea*) *The Odyssey with an English translation by A. T. Murray*. (Loeb Classical Library) London: W. Heinemann, 1919.

Hotelling, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 1933, *24*, 417-441, 498-520.

Houston, C., & Otto, W. Poor readers' functioning on the WISC, Slossen Intelligence Scale, and Quick Test. *Journal of Educational Research*, 1969, *32*, 443-482.

Jöreskog, K. G. Some contributions to maximum likelihood factor analysis. *Psychometrika*, 1967, *32*, 443-482.

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, *23*, 187-200.

Kintsch, W., & Buschke, H. Homophones and synonyms in short-term memory. *Journal of Experimental Psychology*, 1969, *80*, 403-407.

Lawley, D. N. & Maxwell, A. E. *Factor analysis as a statistical method*. London: Butterworths, 1963.

Nabokov, Vladimir. *Bend sinister*. New York: H. Holt, 1947.

Pemberton, C. The closure factors related to other cognitive processes. *Psychometrika*, 1952, *17*, 267-288.

Plutarchus. (*Plutarch's*) *Lives, with an English translation by Bernadotte Perrin*. (Loeb Classical Library) London: W. Heinemann, 1928.

Stanley, J. C., & Wang, M. *The differential weighting of test items*. Princeton, New Jersey: Educational Testing Service, 1969.

Thorndike, E. L., & Lorge, I. A *semantic count of English words*. New York: Institute of Educational Research, 1938.

Thurstone, L. L. Primary mental abilities. *Psychometric Monographs No. 1*, 1938.

Webster's Seventh New Collegiate Dictionary. Springfield, Massachusetts: G. & C. Merriam Company, 1963.

Wetterstrand, W. H. *OBLIQUE: Non orthogonal rotation of factor matrices using the methods of Harris and Kaiser*. University of Wisconsin Computing Center supplementary program series No. 534, 1968.

Wetterstrand, W. H. *FACTOR 1: Factor analysis*. Statjob statistical programs users' manual volume VI. University of Wisconsin Computing Center, 1969.

Winer, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

APPENDIX A

Appendix A

A LIST OF ENGLISH HOMOPHONES

Listed below in alphabetical order are more than 1700 primary sets of English homophones, comprising more than 3800 words. An additional c.1600 secondary (derived) sets, comprising about 3500 words, are indicated by the numerical symbols following some of the primary sets. The list contains, then, a total of about 3300 sets of homophones, with 7300 elements, and includes most of the more commonly encountered English homophones. It is likely that a more comprehensive list could be extended beyond 10,000 elements without exhausting the store of words different in at least one spelling and corresponding meaning, but identical in at least one pronunciation, according to at least one dictionary. The majority of homophones not listed below probably involve: (1) words from the technical vocabularies of commerce, arts, and sciences; (2) words transliterated from modern foreign languages, usually denoting aspects of culture or natural history for which no native English word exists; and (3) words compounded from homophones listed here, since lexicons do not attempt to list exhaustively the compounds of the common prefixes.

Key to the symbols used below

Comma is used to separate successive elements of the homophone sets. Example: yahoo, yahu. (This set contains two elements, YAHOO and YAHU.)

Parentheses are used to indicate alternate spellings (although not all spelling variants are listed). Example: shiek(h). (This word can be spelled SHIEK or SHIEKH without changing the meaning or the pronunciation.)

Hyphen is used to indicate that a word begins with the same string of letters with which the preceding word begins, up to but not necessarily including the letters corresponding to the phoneme represented after the hyphen. Example: aberrance, -ts. (This set consists of ABERRANCE and ABERRANTS.)

Secondary sets are indicated by numerals and underscoring.

The numeral 2 indicates that secondary sets are formed by the addition of -s, -es, or some other plural ending, such as -i or -ae. Example: cast, caste 2. (CASTS and CASTES are also homophones, formed by the 3rd person singular indicative verb suffix and the noun plural suffix.)

The numeral 3 indicates that secondary sets are formed by the addition of the verb endings found in the conjugation of English verbs. These include -(e)s, -(e)d, -ing, as well as the "strong" past tense or past participle of some words, the -est, -eth archaic second and third person singular endings, and the plural form of the gerund. Example: forego, forgo 3. (Also homophones are FOREGOES, FORGOES; FOREWENT, FORWENT; FOREGONE, FORGONE; FOREGOING, FORGOING; FOREGOINGS, FORGOINGS; FOREGOEST, FORGOEST; FOREGOETH, FORGOETH; as well as the periphrastic forms, if one wishes to enumerate them as separate homophone sets. In estimating the number of secondary sets, the symbol 3 was taken as indicating 6.0 additional sets.)

The numeral 4 indicates that secondary sets are formed by the addition of the comparative and superlative suffixes -er and -est. Example: knotty, naughty 4. (Also homophones are KNOTTIER, NAUGHTIER; KNOTTIEST, NAUGHTIEST.)

The numeral 5 indicates that secondary sets are formed by the addition of the suffix -y or -ly (which in the case of monosyllabic or disyllabic stems may in turn take -er and -est suffices). Example: rheum, room 2, 5. (In addition to the plural RHEUMS, ROOMS, indicates by the 2, this set forms the set RHEUMY, ROOMY [and RHEUMIER, ROOMIER; RHEUMIEST, ROOMIEST]).

Since abridged dictionaries frequently neglect the infrequent use of adjectives as nouns, nouns as verbs, or the formation of adjectives or adverbs in -ly from simpler words, it is likely that there are many secondary sets which can be formed from sets listed below which are not indicated by the numerical code.

If only a subset of a homophone set forms secondary sets, that subset is underscored, and the code number of the secondary set is correspondingly underscored. Example: vary, verry, very 2, 4. (The secondary sets VARIES, VERRIES, and VARIER, VERIER; VARIEST, VERIEST can be formed from this set.) The use of both solid and broken lines to indicate secondary sets below is quite infrequent.

If the code numbers of secondary sets of the same primary set are underscored by n and (n + 1) solid lines, for n = 0 or 1, then words with (n + 1) underlines are also generators of secondary sets coded by the symbol with n lines. Example: vain, vane, vein 2, 4.

(All three words form a secondary set in -s, VAINS, VANES, VEINS, since 2 has 0 underlines; while VAINER, VEINER; VAINEST, VEINEST are secondary sets formed by the underlined rule on the underlined words.)

An asterisk following a primary set indicates that all of the secondary sets (usually those in -s or -es) derived from the set are homophones of some uninflected word; this symbol is used as an aid to those who may wish to expand the list by adding secondary sets neglected by the present investigator. Example: crew, krew *. (The plural forms of these words have homophones CROUSE, CRUISE, CRUS, and CRUSE. The word KREW(E), unlisted by many dictionaries, denotes a marching group in a New Orleans Mardi Gras parade.)

Most of the words that follow have been classified as homophones on the authority of one or more of the following sources:

__________, *Webster's New Seventh Collegiate Dictionary*. Springfield, Massachusetts: G. & C. Merriam Company, 1963.

Barnhart, C. L. (Editor), *The American College Dictionary*. New York: Random House, 1960.

Bey, A. F. Inglott, *A Dictionary of English Homonyms, Pronouncing and Explanatory*. London: K. Paul, Trench, Trubner, 1899.

Franklyn, Julian, *Which Witch? being a grouping of phonetically compatible words*. London: H. Hamilton, 1966.

Stein, Jess. *The Random House Dictionary of the English Language* (Unabridged edition). New York: Random House, 1966.

a, eh
aberrance, -ts
aboard, abord
abrade, abraid, abrayed 3
absence, -ts
accidence, -ts
accur, occur 3
acetic, ascetic
acher, achor, acker, acre, akre 2
acts, ax (e)
ad, add *
adds, ads, adze
adeps, adepts
adieu, ado 2
adherence, -ts
adolescence, -ts
adulteress, adulterous
adventuress, adventurous
aerial, ariel, areal 2
aerie, airy
aerie, eerie
affiance, -ts
affluence, -ts
affrayed, afraid
ah, I

ai, ay, aye, eye, I 2, 3

aid, aide 2

ail, ale 2

air, aire, are, ayre, e'er, ere, err, eyre, heir 2, 3

aire, ara

airship, heirship 2

aisle, I'll, isle 2

ait, ate, eight *

aits, eights, eighths

alderman, aldermen, ealdorman, ealdormen 2

alkalis, alkalize

all, awl 2

allowed, aloud

altar, alter 2

amine, ammine 2

ana, anna 2

analyst, annalist 2

anele, anneal 3

anger, angur 3

ant, aunt 2

antae, ante, anti, aunty 2

antecedence, -ts

apatite, appetite 2

apothegm, apothem 2

appose, oppose 3

apposition, opposition 2

appositive, oppositive 2

appressed, oppressed

appurtenance, -ts

ar, are, err, or, our 2

arc, ark 2

area, aria 2

aril, aryl 2

arris, heiress 2

ascendance, -ts

ascent, assent 2

assistance, -ts

assonance, -ts

attendance, -ts

aught, ought

aura, ora 2

aural, oral

aureole, oriole 2

aurical, oracle 2

aw, awe

awful, offal

awn, on

B, be, bee *

ba, baa, bah 2

bac, back

bacchant, -e 2

bach, batch 2

bacilary, basilary

bad, bade

bade, bayed

baetyl, beadle, Beatle, beetle, betel

bail, bale, Baal 3, 4

bailee, bailey, bailie 2

bait, bate 3

baize, bays, beys

bald, balled, bawled

balk, bock *

balks, bocks, box

ball, bawl 3

balm, bomb, bombe 2

band, banned

bands, banns, bans

bang, bhang 2

banian, banyan 2

bar, barre 2

bard(e), barred

bare, bear 3, 4

bark, barque 2

baron, barren 2

barrier, burier 2

basal, basil 2

base, bass, beth 2, 4

based, baste

basest, bassist

bask, basque, Basque 2

baton, batten 2

bauble, bobble 2

bay, bey *

bazaar, bizarre 2

beach, beech 2

beading, beating 2
bean, been

beard, beered

beat, beet 2

beau, bo, boe, bow 2

beaut, butte 2

been, bin

beer, bier, birr 2

bees, bise, B's

bel, bell, belle 2

belligerence, -ts

benzene, benzine

berg, bourg, burg 2

berry, bury 3

berth, birth 3

better, bettor 2

bided, bighted

biding, bighting, biting 2

bight, bite 3

billed, build

binds, bines

binnacle, binocle 2

bird, burd, burred 2

birl, burl 3

birn, burn 2

bit, bitt 2

blend, -e 2

blew, blue

bloc, block 2

blowze, blouse 2

boar, bore 2

board, bored

boarder, border 2

boating, boding 2

bode, bowed

bold, bowled

bolder, boulder

bole, boll, bowl 2

bonds, bonze

boos, booze

booty, bootie (-tee) 2.

born, borne, bourne

borough, borro, burrow 2

bough, bow *

bougnts, bows, bowse (-ouse)

boule, boulle (buhl) 2

boullion, bullion 2

bourden, burden 2

bourse, burse 2

boy, buoy 2

bra, braw

brache, brash 2

brachs, bracts

brae, bray, brey 3 *

braes, braise, brays, braze, breys 3

braid, brayed, breyed

brail, braille 3

brands, brans

brassie, brassy

breach, breech 2

bread, bred

breathes, brees, breeze

breed, brede 2

brewed, brood

brews, bruise

bridal, bridle 2

bricks, brix

brilliance, -ts

broach, brooch 2

brooded, bruited

brooding, bruiting 2

broom, brume 2
brows, browse
bruit, brut, brute 2
brunet, brunette 2, 4
brunetness, brunetteness
buccal, buckle
budded, butted
budding, butting 2
bullae, bully
bunds, buns
burger, burgher 2
burley, burly
bus, buss
bussed, bust
but, butt 2
buy, by, bye 2
buyer, byre 2

C, cee, sea, see, si *
cache, cash 3
cachou, cashew 2
calendar, calender 3
calix, calyx 2
calk, caulk 3 *
calks, calx, caulks
call, caul 2
callous, callus

cam, cham 2

camail, camaille

came, kame 2

campaign, campane

can, khan 2

canape, canopy 2

canned, kand

cannon, canon 2

can't, cant, quinte 2

canter, cantor 2

capital, capitol 2

caracul, karakul 2

caret, carrot, karat 2

caries, carries

carol, carrel 2

carpal, carpel 2

cart, carte, quarte 2

cask, casque 2

cast, caste 2

caster, castor 2

caudal, coddle 2

cause, caws

cay, K, quai (quay) 2

cedar, ceder, seater, seder, seeder 2

cede, seed 3

ceded, seated, seeded

ceding, seating, seeding 2

cees, C's, seas, sees, seise, seize, sis

ceil, seal, seel 3

cell, sell, selle 2

cellar, seller 2

cense, cents, scents, sense 3

censer, censor, sensor 2

censorial, sensorial 5

cent, scent, sent *

cerate, cirrate

cere, sear, ser, sere, seer 2, 3

cereal, serial 2

Ceres, series

cereous, serious, Sirius

ceric, xeric

cero, serow 2

cerous, cirrous, cirrus, scirrhous, scirrhus, seeress, serous 2, 2 (-i)

cession, session 2

cetaceous, setaceous

chair, chare 3

chaise, chays (shays), sheas

champagne, champaign 2

chance, chants

chantey, shanty 2

chard, charred

chased, chaste

chatelain, chatelaine 2
chatoyance, -ts
chauffeur, shofar 2
cheap, cheep 4
check, cheque, Czech 2
checker, chequer 2
chews, choose
chic, shiek (h) 2
chilly, chili (chile, chilli)
chitin, chiton 2
choir, quire 2
cholate, collate 2
choler, collar 2
choler, coaler 2
choline, colleen 2
choral(e), coral 2
choral(e), corral 2
chord, cord, cored 2
chordate, cordate
choria, coria
chose, shows
chott, shot 2
chrism, chrisom 2
chute, shoot 3
cilia, psyllia
cinque, sink, sync 2, 3

cirrhosis, sorosis 2
cist, kissed
cist, cyst 2
cite, sight, site 2, 3
cited, sided, sighted
citeless, sightless 5
citing, siding, sighting 2
citrin, citrine, citron
clack, claque 2
claimant, clamant
clair, clare
clamber, clamor 3
claus, clause 2
clefs, clefts
climb, clime 2
clique, cleek 2
clique, click 2
close, clothes, cloze 2
coal, cole, kohl *
coaled, cold
coals, colds, coles, kohls
coaming, combing 2
coarse, corse, course 2, 4
coat, cote 2
coating, coding 2
coax, cokes
cob, cobb, kob 2
cobble, coble 2

coccal, cockle 2

cocks, cox

coco(a), koko 2

coela, selah

coffer, cougher 2

coif, quaff 3

coign, coin, quoin 3

coincidence, -ts

coir, coyer

cola, kola 2

collard, collared

collie, colly 2

cologne, colon 2

colonel, kernel 2

color, culler 2

comb, combe 2

come, cum

complement, compliment

complementary, complimentary 5

compt, count 2

comptroller, controller 2

con, khan 2

concent, consent 2

concenter, consenter 2

concomitance, -ts

concurrence, -ts

confidance, -tes, -ts

confidant, -e *

confluence, -ts

consequence, -ts

consonance, -ts

continence, -ts

contingence, -ts

continuance, -ts

convalescence, -ts

coo, coup 2

coolie, coolly, coulee 2

coop, coup, coupe 2

cop, kop *

cops, copse, kops

copy, kopje 2

core, corps, kor 2

corollate, correlate

correspondence, -ts

cosign, cosine 2

cot, cote 2

cough, kaph 2

could, cud

council, counsel 2

councilor (-er), counselor (-er) 2

councilship, counselship 2

courier, currier 2

cousin, cozen 2

covariance, -ts

coward, cowered
craft, kraft 2
crater, krater 2
crape, crepe 2
creak, creek 2
cream, creme 2
crease, kris 2
crew, krew *
crewed, crowd, crude
crewel, cruel
crews, crouse, cruise, crus, cruse, krews 2
crooks, crux
croon, kroon 2
cross, crosse
crows, croze
crumby, crummy 4
cue, Q, queue 2, 3
cumulous, cumulus
cur, curr 2
curb, kerb 2
curdle, curtal, kirtle 2
currant, current 2
curser, cursor 2
cyanite, syenite 2
cygnet, signet 2
cylix, siliques
cymas, cymous
cymbal, symbol 2

cymbalist, symbolist 2

cynicism, sinicism 2

cypress, cyprus 2

cystine, sistine

cytology, sitology 2

D, de, di 2

dace, dais

daffed, daft

dairy, derry 2

dais, deus

dam, damn 3

dandie, dandy

daughter, dodder, dotter 2

dawn, don 3

day, dey *

days, daze, deys

deal, deil 2

dean, dene 2

dear, deer 2

deasil, diesel 2

debutant, -e 2

decadence, -ts

decern, discern 3

decks, dex

deem, deme 2

deference, -ts
del, dell 2
delate, dilate 3
delation, dilation 2
delative, dilative 2
delator, dilator 2
demean, demesne 2
dental, dentil
dense, dents
dependance (-ence), dependants (-ents)
descendant, descendent
descent, dissent 2
desert, dessert 2
deterrence, -ts
deuce, douce, douse 2
deuced, doused
deviance, -ts
deviser, devisor, divisor 2
dew, do, doo, doux, due 2, 3
dewed, dude
dextran, dextrin(e) 2
dhikr, dicker 2
dhole, dole 2
dhoope, dupe 2
dhootie, duty 2
dhou, dow, tao*

dhous, dows, dowse

die, dye 3 *

diesis, diocese

dike, dyke 2

diker, duiker 2

dine, dyne 2

dinest, dynast

dinghy, dinky 2

dire, dyer

discreet, discrete 4, 5

discreetness, discreteness 2

dissidence, -ts

divorce, -ee 2

do, doe, dough, dow *

docile, dossil

does, dos, doughts, dows, doze

dolman, dolmen

dominance, -ts

domine, dominie

dominoes, dominos

done, dun

donjon, dungeon 2

d(h)ooly, duly

door, dor 2

dope, dowp 2

dost, dust

doubt, dought

doubter, douter 2

dour, dower

dour, dure

douse, dowse 2

doyen, doyenne 2

drachm, dram 2

draffs, drafts, draughts

draft, draught *

droop, drupe 2

dual, duel 2

dualist, duel(l)ist 2

ducked, duct

ducks, ducts

dura, durra (doura[h]) 2

eager, eagre

eagle, egal

earing, earring 2

earn, ern(e), urn 2

ease, E's

eau, eaux, O, oh, owe 2

eaves, eves

een, e'en, eyen (eyne)

efference, -ts

efts, F's

el, ell, L 2

elision, elysian

elusion, illusion 2

elusive, illusive

elusiveness, illusiveness 2

elusory, illusory

em, 'em, M 2

embracer, embraceor 2

emerge, immerge 3

emergence, -ts

encyst, insist 3

ends, ens

entrance, -ts

equivalence, -ts

ere, err, eyre 2

erred, urd

errupt, irrupt 3

ethel, ethyl 2

ewe, U, yew, you *

ewer, your, you're *

ewes, U's, use, yews, yous(e), youths

ewers, yours

exert, exsert 3

expedience, -ts

faddist, fattest

faded, fated

faem, fame 2

faerie, fairy, ferry 2

faille, file, phial 2

fain, fane, feign, foehn 2, 4

faint, feint 3, 4

fair, fare 3, 4

faker, fakir 2

faro, farrow, pharaoh 2

fat, phat 4

fate, fete 2

faun, fawn

fay, fey *

fays, faze, fease (feeze), phase 2, 3

feal, feel 3

fear, fere 2

feat, feet, fete

feaze (feeze), fees, phis

fee, phi *

feels, fields

felloe, fellow 2

felt, vel(d)t 2

feme, femme 2

fends, fens

fennel, phenyl 2

feoff, fief 2

feral, ferrule, ferule 2

fern, foehn, foin 2

feudal, futile

few, phew

fiance, fiancee 2

fie, phi 2

fil, fill

filar, phylar

file, phial 2

filter, philter (philtre) 2

filum, phylum 2(-a), 2(-s)

find, fined

finds, fines

fir, fur *

firry, furry 4

firs, furs, furze

fisher, fissure 2

fizz, phiz 2

flack, flak *

flacks, flaks, flax

flair, flare 2

flamboyance, -ts

flautest, flautist

flea, flee 2

fleche, flesh 2

flecks, flex

flew, flue, flu

flight, flite 3

floc, flock *

flocks, flocs, phlox

floe, flow 2

floresce, fluoresce 3

florescence, -ts, fluorescence, -ts 2

florescent, fluorescent *

flour, flower 3, 5

foaled, fold

for, fore, 'fore, four 2

foram, forum

forbear, forebear 2

force, fourths

forcene, foreseen

forefend, forfend 3

forego, forgo 3 (-went, -gone)

forehanded, fourhanded

forelay, forlay 3

foreword, forward

forewords, forwards

fort, forte 2

forte, forty *

fortes, forties

forth, fourth

foul, fowl 2

fraise, frays, phrase 2

franc, frank 2

frap, frappe 2

frees, freeze, frieze 2

friar, fryer (frier) 2

fro(w), froe, vrau *

froes (frows), froze, vraus

fugal, fugle

fund, funned

funds, funs

fungous, fungus

furor, furore 2

fusel, fusil(e) 2

gaff, gaffe 2

gage, guage 2

gait, gate 2

gaited, gated

gallein, galleon 2

galley, gally 2

gallop, galop 2

gallous, gallus

gamble, gambol 3

gang, gangue 2

gassed, ghast

gaud, god 2

gel, jell 2

genap, genappe

genes, jeans

genet, jennet 2

german, germen

geste, jest 2

ghat, got

ghoul, gul 2

gibber, jibber 2

gibe, jibe 3

gild, gilled, guild 2

gilder, guilder 2

gilt, guilt 2

gimel, gimmal 2

gin, (d)jinn(h)

glair(e), glare 3

glair(e)y, glarey 4

glairiness, glariness 2

glands, glans

glaze, gleys

glean, gleen

gled(e), gleed 2

glossae, glossy

glows, gloze

gloom, glume 2

gnar(r)ed, nard

gnawer, nor

gneisse, nice

gnome, nome 2

gnome, gnomey

gnu, knew, new, nu *

gnus, news, nus

goaled, gold

goals, golds

goer, gore 2

goffer, golfer 2

gook, guck

gored, gourd

gorilla, guerilla 2

grade, grayed (greyed)

grader, grater, greater 2

grading, grating 2

grafts, graphs

graft, graphed

graip, grape 2

grate, great *

gray (grey), gres

grays (greys), graze

grease, gr(i)ese, gris 2

grease, grees

greave, grieve 2

grew, gru(e) 2

grievance, -ts

griff, griffe *

griffes, griffs, grifts

grill, grille 2

grip, grippe 2

grisly, grizzly 4

groan, grown

grocer, grosser 2

groin, groyne 2

guide, guyed

guise, guys

gunnel, gunwale 2

gyve, jive 3

hade, hayed

hading, hating 2

haem, hem 2

haik, hike 2

hail, hale 2

hair, hare 2

handmade, handmaid 2

hangar, hanger 2

handsome, hansom

hairy, harry 4

hart, heart

hall, haul, hawl 2, 3

halve, have 3

have, of

haws, hawse

hay, he, heigh, hey *

hays, haze, heighs, hes, heys

heal, heel, he'll, hill, 2, 3, 4

hear, here

heard, herd

heating, heeding 2

he'd, heed

height, hight

heigh, hi, hie, high 2

hemidactylous, -lus

hence, hents

herl, hurl 2

hertz, hurts

hew, hue, whew 2

hewed, hued

hic, hick

hide, hied

higher, hire

hike, yoicks

him, hymn

his, is

hissed, hist

ho, hoe *

hoar, hoer, whore 2, 4

hoard, horde, whored 2

hoarse, horse

hoarsen, whoreson 2

hold, holed

holds, holes, wholes

hole, whole 2

hollo, hollow 2

holm, home 2

holey, holy, wholly 4

homage, ohmage 2

hombre, ombre 2

hos, hoes, hose

hostel, hostile

hour, our 2

house, hows

humerous, humerus

humeral, humoral

humous, humus

hydrocele, hydrocoel(e) 2

hydrocephalous, -lus

idle, idol, idyll 2

illuminance, -ts

immanence, -ts, imminence 2

immanent, imminent

impotence, -ts

impressed, imprest

in, inn 2

incidence, -ts

inciter, insider 2

indict, indite 3, 4

indiscreet, indiscrete 2, 5

influence, -ts

innocence, -ts

insolant, insulant *

insolants, insolence, insulants

instance, -ts

insurgence, -ts

intendance, -ts

intense, -ts

intension, intention 3

intercession, intersession 2

intern, interne 2

interpellation, interpolation 2

intersects, intersex

intransigence, -ts

invade, inveighed

invariance -ts

its, it's

jam, jamb 2
jaunce, -ts
jewel, joule 2
jinks, jinx
junkie, junky
just, gist 2
just, jest 2
just, joust 2, 4
just, jus

kail, kale 2
kayak, kyack 2
key, ki 2
kill, kiln 2
kite, kyte 2
kittel, kittle 2
knacker, nacre 2
knap, nap, nape, nappe 2, 3
knave, nave 2
knead, kneed, need 3
knickers, nickers
knight, night 2, 5
knit, nit 2
knitted, nited
knob, nob 3
knock, nock 3
knot, naught, not, nought 2
knotty, naughty 4

knout, newt 2

know, no, noh *

knowns, nones

knows, nohs, nos, nose

krona, krone

la, law

label, labile

lac, lack, lakh *

lache, lash

laches, latches

lacks, lacs, lakhs, lax

ladder, latter

lade, laid, layed

lager, logger 2

lain, lane

lair, layer 2

lam, lamb 3

lama, llama 2

lance, launce 2

laps, lapse

laud, lawed

laurel, loral

lay, lea, lei, les, ley *

lays, laze, leas, leis, leys

lea, lee, ley, li 2

leach, leech 3

lead, led

lead, lied

leader, lieder, liter (litre) 2

leaf, lief

leak, leek 2

lean, lene, lien 2

leary, leery 4

leased, least

leave, lief

leaven, levin

leks, lex

lends, lens

less, loess

lessen, lesson 2

lets, let's

levee, levy 3

lever, liever, livre 2

lewd, lood

lewder, looter

lewdest, lutist

liar, lier, lyre 2

liable, libel

lice, lyse

lichen, liken 2

lie, lye *

lies, lyes, lyse

lighten, lightning 2

limb, limn 2

limey, limy

links, lynx

literal, littoral

lo, low 2

load, lode, lowed 2

loan, lone 4

loath, loathe

loch, lock, loughs *

lochs, locks, loughs, lox

looks, luxe

loons, lunes

loop, loupe 2

loos, lose

loose, luce 2

loot, lute 3

lorry, lory 2

louie, louis *

louies, louis, lues

lucks, lux, luxe

lumbar, lumber 2

lumen, lumine

luster, lustre 2

lysin, lysine 2

ma, maw 2

mach, mock 2

madam, -e 2

madding, matting

made, maid

magnate, magnet 2

mail, male 2

mailer, malar 2

main, mane 2

maitre, meter 2

maize, mays, maze 2

malign, moline

mall, maul, moll 2

ma'am, mom 2

mammae, mammy

mandrel (mandril), mandrill 2

manes, mainies

manner, manor 2

man(n)iken, mannequin 2

mantel, mantle 2

marc, mark, marque 2

mare, mayor 2

marine, moreen 2

marlin, marline 2

marquee, marquis *

marquees, marquis, marquise, marquises

marquis, marquise 2

marry, merry

marshal, martial

marten, martin 2

maser, mazer 2

mask, masque 3

masker, masquer 2

massed, mast

mat, matte 2

me, mi

mead, meed 2

meak, meek

mean, mesne, mien 2

meanings, meninx

meat, meet, mete 3

meatier, meteor

medal, meddle, metal, mettle 2, 3

meddler, medlar 2

meeter, meter, metre 2

mere, mirror 2

merlin, merlon 2

mesic, tmesic

metis, metisse, metisses

mew, mu *

mewl, mule 2

mews, mus, muse
mho, mot, mow *
mhos, mohs, mots, mows
might, mite 2
mil, mill 2
milch, milk
millenary, millinery 2
millrace, milreis
mina, myna(h) 2
mince, mints
mind, mined
minds, mines
miner, minor 2
minks, minx
mis, mise
missal, missel, missile 2
missed, mist
misses, missis (missus), Mrs.
moan, mown
moat, mote 2
mode, mowed
moire, mohur, mor, more, mower 2
moire, moray *
moires, morays, mores
mol, mole *
molten, moulten

mo(u)lds, moles, mols

moo, moue, mu 2

mood, mooed

moola(h), mullah

moose, mousse 2

mordant, mordent

morn, mourn 2

morning, mourning 2

morro, morrow 2

mouse, mows

mousseline, mousselines

mucous, mucus

mudder, mutter 2

mum, mumm 2

murderess, murderous

murre, myrrh 2

muscle, mussel 2

mussed, must

mustard, mustered

mustee, musty

mutual, mutuel 2

nae, nay, ne, nee, neigh *

naval, navel

nays, naze, neighs

necklace, neckless

nicks, nix, pnyx
nil, nill
nocturn, nocturne
none, nun
nonabsorbence, -ts
nonadherence, -ts
noncereal, nonserial
nonconcurrence, -ts
nonresidence, -ts
noose, nous 2
novelest, novelist
novocain, novocaine 2
nudest, nudist
nymphs, nymss

oak, oke
oar, o'er, or, ore 3
observance, -ts
occellated, oscillated
occellation, oscillation 2
occident, oxidant 2
od, odd 2
od, ode, owed 2
odder, otter
odeum, odium 2
(o)estrous, (o)estrus

offed, oft

ohm, om (aum) 2

oleo, olio 2

ombudsman, -men

once, wonts

one, won

oohs, ooze

orthopteran, -on

outcast, -caste

outright, -write

outrode, -rowed

outwait, outweight 3

overbilled, overbuild

overdo, overdue

overhigher, overhire

overlade, overlaid

overman, overmen

oversea, oversee 2

oxeyed, oxide

P, pea, pee *

pa, pah

pace, pes

paced, paste

packed, pact

packs, pacts, pax (Pax)

paean, paeon 2

paeon, peon 2

paid, payed

pail, pale 2

pain, pane 2

pair, pare, pear 2, 3, 4

palate, palette, pallet, pallette 2

palette knife, pallette knife

pall, pawl 2

palmar, palmer

pan, panne 2

papous, pappus

par, parr *

pard, parred

pardner, pardoner 2

parish, perish 2

parlay, parley 3

parol, parole 2

parrs, pars, parse

passed, past

patchily, patchouly

paten, patten 2

pater, patter 2

patience, -ts

pause, paws

pawed, pod

pawned, pond

pawns, ponds, pons

pay, pe *

pays, peise, pes

peace, piece 2

peak, peek, pique 3

peal, peel 3

pearl, purl 3

peas(e), pees, peise, P's

pecten, pectin 2

pedal, peddle, petal 2, 3

peer, pier 2

pelisse, police 2

pekoe, picot 2

penance, pennants

pencel, pencil, pensile 2

pendant, pendent 2

penetrance, -ts

penitence, -ts

penni, penny

peony, piney

people, pipal 2

pepless, peplos (peplus)

per, purr

permanence, -ts

pervade, purveyed

petrel, petrol 2

phew, whew

phosphene, phosphine 2

phylae, phyle

pi, pie, (pye) 2

pic, pick *

pica, pika 2

pickle, picul 2

picks, pics, pix, pyx

picnic, pyknic 2

pidgin, pigeon 2

pimento, pimiento 2

pincer, pincher 2

pique, piquet

pistil, pistol 2

pixy, pyxie

place, plaice

plain, plane 3, 4

plainer, planar, planer

plait, plate 3

plait, plat 3

plaiter, plater 2

plantar, planter

platan, platen 2

pleas, please

pleural, plural

plodding, plotting 2

plum, plumb 2

pocks, pox

pocus, polkas

poem, pome 2

pogey, pogy 2

pois, poise

polar, poler, poller 2

pole, poll 3

policies, pollices

policlinic, polyclinic 2

politic, politick 2

pollex, pollocks (-llacks), Pollux

pomace, pumice 2

pommel, pummel 3

ponds, pons

pood, pooed

pool, pul 2

poor, pore, pour 3

poorer, porer, pourer 2

populace, populous

pose, pows

practice, practise 2

praise, prase, prays, preys 2

prao, prow 2

pray, prey 3

prayer, preyer 2
preadolescence, -ts
precedence, -ts
precisian, precision 2
precisianism, precisionism 2
premier, première 2
presence, -ts
pressed, prest
presser, pressor 2
prevalence, -ts
pride, pried
prier (pryer), prior 2
pries, prize (prise)
primer, primmer
prince, prints
princess, princesse
principal, principle 2
profit, prophet 2
propellant, propellent
pros, prose
protean, protein
protege, protegee 2
proud, prowed
psalter, salter 2
psammite, samite 2
psephite, seafight 2

pshaw, shah, shaw 2
psi, sigh, xi *
psis, sighs, size, xis
psych, sike 2
psychosis, sycosis 2
psylla, scilla, scylla 2
puisne, puny
puli, pulley
pupal, pupil
pupilary, pupillary
puree, purey
pursy, pussy 4
puttee, putty 2
pyric, pyrrhic

quad, quod 2
quaere, query 2
quaestor, questor 2
quarts, quartz
quean, queen 2
queerest, querest
quince, -ts
quintal, quintile 2
quitter, quittor 2

rabbet, rabbit 3
race, reis, res
rack, wrack 3
rackets, racquets
raddle, rattle 3
radiance, -ts
radical, radicle 2
radish, reddish
raffs, rafts
ragman, ragmen
rah, raw
raider, rater 2
raiding, rating 2
rain, reign, rein 3
raise, rays, raze (rase), res 3
raiser, razer (raser), razor 2
rami, ramie
ramous, ramus
rancor, ranker 2
rap, wrap 3
rapped, rapt, wrapped (wrapt)
rapper, wrapper 2
ratal, ratel 2
rath, wrath
rathe, wraith
raven, ravin 2
rawer, roar

ray, re*

re, ree 2

reactance, -ts

read, red, redd 2

read, rede, reed 3

reader, reder, rhetor 2

real, reel, riel 2, 4

real, rial, riyal 2

ream, reem, riem 2

reave, reeve, reive 3

reaver, reever, reiver 2

rebait, rebate 3

rebilled, rebuild

recede, reseed 3

receded, receipted, reseated, reseeded

receding, receipting, reseating, reseeding 2

receipt, reseat 3

recipience, -ts

recision, rescission 2

reck, wreck 3

recks, rex, wrecks

redocks, redox

reek, wreak 3

reference, -ts

reffed, reft

refind, refined

reflection, reflexion 2
reflects, reflex
regard, reguard 3
reheal, reheel 3
remark, remarque 2
remonstrance, -ts
renin, rennin 2
repassed, repast
reroot, reroute 3
resail, resale 2
residence, -ts
resistance, -ts
resold, resoled
resonance, -ts
respondence, -ts
rest, wrest 3
rester, wrester 2
retard, retarred
retch, wretch 2
retracked, retract
revere, revers
review, revue 2
rheum, room 2, 5
rho, roe, row *
rhos, roes, rose, rows
rhumb, rum 2
rhyme, rime 3

riband, ribboned
ribands, ribbons
ride, wried
rider, righter, writer 2
riding, righting, writing 2
riffed, rift
riffs, rifts
rigger, rigor 2
right, rite, wright, write 2, 3
ring, wring 3
ringer, wringer 2
riot, ryat 2
rise, ryes, wries
road, rode, rowed
roan, rowan 2
robands (robbins), robins
robbin, robin *
roc, rock 2
roil, royal 2, 5
role, roll 2
rondeau, rondo 2
rood, rude, rued
rookie, rooky
roomer, rumor 2
roose, rues, ruse 2
root, rout, route 3
rooter, router, ruder 2

rot, wrought
rote, wrote
rough, ruff 3
rougher, ruffer 2
roupy, rupee
rouse, rows
roux, rue
rowan, rowen 2
rubble, ruble 2
rudder, rutter 2
rye, wry *

sac, sack, sacque *
saccharin, -e
sachet, sashay 2
sacks, sacques, sacs, sax
saddest, sadist
sadhe, soddy
sail, sale 2
sailer, sailor 2
sain, sane, seine 3
salience, -ts
salver, salvor 2
same, sejm
sandhi, sandy
sands, sans

sari, sorry
satiric, satyric 5
satirical, satyrical 5
sault, Sioux, sou, sue 2
saurel, sorrel
saury, sori, sorry
saver, savor 2
sawer, soar, sore 2
scald, skald 2
scalar, scaler 2
scat, scot 2
scat, skat 2
scat, scot 2
scat, skat 2
scar(e)y, skerry
scend, send 3
scends, sends, sens
scene, seen, sin
scenery, senary
schiller, shiller 2
schilling, shilling, skilling 2
scirrhous, scirrhus
scold, skoaled
scolds, skoals
scoot, scute 2
scop, scope 2

scop, shop 2
scrips, scripts
scrooge, scrouge 2
scull, skull 3
scurry, skerry 2
seaborn, seaborne
seam, seem, seme,xeme 2, 3
seaman, seamen, semen
season, seizin (seisin) 2
second, secund 5
sects, sex
seek, sic, sikh 2, 3
seidel, sidle 2
seigneury, seignory 2
seignoir, senior 2
seised, seized
semens, siemens
senate, sennet, sennit 2
sequence, -ts
sera, sirrah
seraph, serif 2
serf, surf 2
serge, surge 2
serin, serine
serval, servile
settler, settlor 2
sew, so, sow 3

sewer, soar, sore, sower 2

sewer, suer 2

sewn, sone, sown

sexed, sext

shake, shiek(h) 2

shammes, shamus

shay, shea *

she, shea, Sidhe, ski 2

shear, sheer, skier, 2, 3

sheen, shin 2

she'll, shill

sherd, shirred

shier (shyer), shire 2

shirr, sure

shoe, shoo 3

shone, shown (shewn)

shore, shower 2

shrewd, shrewed

sibilance, -ts

sic, sick *

sicks, sics, six, sixth

side, sighed

sigher, sire 2

sign, sine, syne 2

significance, -ts

siphon, syphon 2

sirred, surd

sixte, sixth 2
sizar, sizer 2
skeet, skete 2
slay, sleigh 3
slayer, sleigher 2
sleave, sleeve 3
sleight, slight 2
slew, slough, slue 3
sliding, slighting 2
slipe, slype 2
sloe, slow 2
soak, soke 2
soared, sword
softs, sophs
sol, sole, soul 2
solace, solus, soulless
sold, soled, souled
some, sum
son, sun, sunn 2
sonance, -ts
sonny, sunny
soot, suit, suite 3
sorceress, sorcerous
sough, sow *
soughs, souse, sows

snees, sneeze

spade, spayed

speck, spec

speel, spiel 3

speiss, spice 2

spicae, spic(e)y

spier, spire 2

spinner, spinor 2

spirituel, -elle

spital, spittle 2

spits, spitz

spoor, spore 3

squaller, squalor 2

stade, staid, stayed

staff, staph 2

stain, stane 2

stair, stare 2

stake, steak 2

stanch, staunch

stancher, stauncher

starlet, starlit

stater, stator 2

stationary, stationery 2

steal, steel, stele 2, 3

stealer, steeler, stelar 2

stearic, steric

steely, stele, stelae 2

steer, stere 2

step, steppe 2

stile, style 2

stolen, stollen, stolon 2

stoop, stoup, stupe 2

storey, story 2

stoss, stows

straight, strait 2, 4

straightly, straitly 4

straightness, straitness 2

straighten, straiten 3

streak, streek, 2, 4

strider, stridor 2

strode, strowed

stooper, stupor 2

sty, stye

stylar, styler

suade, suede, swayed 2

subbase, subbass 2

subsequence, -ts

subtile, subtle 4, 5

subtileness, subtleness 2

subtiler, subtler, sutler

succor, sucker 3

succubae, -i

succulence, -ts

sudds, suds

suite, sweet 2

sulci, sulky

sulfur, sulphur 2

summary, summery

sundae, Sunday 2

superintendence, -ts

suppliance, -ts

sura, surah, surra 2

surplice, surplus 2

surveilance, -ts

susurrous, susurrus

swat, swot 3

swooned, swound

swoons, 'swounds

T, tea, tee, ti *

tace (tasse), tass 2

tace, teth 2

tach, tack *

tachs, tacks, tacts, tax

tacit, tasset

tacked, tact

tael, tail, taille, tale 2

ta'en, tain

tahr (thar), tar, tarre 2

tailer, tailor 2

taint, teint

taler (thaler), taller

tamarinds, tammarins

Tao, tau

taper, tapir 2

tare, tear 3

tartar, tarter

tartarous, Tartarus

tau, taw 2

taught, taut

taupe, tope 2

Tauri, tori, torrey, tory

Taurus, torus *

taus, touse

tav (taw), toff *

tavs (taws), toffs, tofts

taxes, taxis

taxis, taxus

teal, teil 2

team, teem 3

tear, tier 2

tearer, terror 2

T's, teas, tease, tees, tis

teat, tit 2

tele, telly 2

tellurian, tellurion 2

temps, tempts

tenace, tennis 2

tends, tens

tennis, tenous

tense, tents, tenths

tenser, tensor

terce, terse, turse

tern, terne, turn 2

ternary, turnery

terrain, terrane 2

terrene, tureen

thallous, thallus

thane, thegn 2

the (ye), thee

their, there, they're

therefor, therefore

thes, these

thirl, thurl 2

thrash, thresh 3

thrasher, thresher 2

threw, through, thru

throe, throw 2

throne, thrown

thyme, time 2

thyrses, thyrsus

tic, tick 2

tical, tickle 2

tide, tied

tie, tye 3

tide, tied, tyed

tier, tire (tyre), tyer 2

tighter, titer (titre)

'til, till

timbal, timbale 2

timber, timbre 2

tincal, tinkle

tip, typp 2

to, too, tu, two

toad, toed, towed

toady, tody 2

tocsin, toxin 2

toe, tow 2

toffed, toft

toise, toys

toke, toque 2

told, toled, tolled

tole, toll 3

ton, tun 2

tongue, tung 2

took, touk
took, tuck
tool, tulle 2
toon, tune 2
toona, tuna 2
toot, tout 3
tooter, tudor, tutor, twodoor 2
tor, tore, torr 2
tort, torte 2
tough, tuff *
toughs, tuffs, tufts
touse, twos
tracked, tract
tracks, tracts
trader, traitor 2
traitorous, traitress
trance, trans, trants
transferer, -or 2
transience, -ts
tray, tres, trey 2
treest, triste
trews, trues
tri, try 2
triaene, triene
trichy, tricky
tricorn, tricorne

trigon, trygon 2
trios, triose
triptych, tryptic
trivet, trivvet 2
troche, trochee, trochi 2
trollop, trollope 2
troolie, truly
troop, troupe 3
trooper, trouper 2
trophy, trophi
trough, trow 2
trussed, trust
trustee, trusty
tubulous, tubulus
tucks, tux
tumulous, tumulus
turkey, turki
tussal, tussle
twees, tweeze
typhous, typhus
tyrannis, tyrrannous

uca, yuca
udder, utter 2
ugli, ugly 2
ulmous, ulmus

umber, umbre 2
unbilled, unbuild
underbilled, underbuild
underhold, underholed
underman, undermen
undo, undue
unfrees, unfreeze
unlade, unlaid
unkneaded, unneeded
unreal, unreel
unwanted, unwonted
unseal, unseel 3
unright, unwrite 2

vacillate, vassalate 3
vail, vale, veil 3
vain, vane, vein 2, 4
valance, valence 2
valiance, -ts
vaned, veined
vary, very, very 2, 4
vau, vow 2
vault, volt 2
vela, vila
versed, verst, wurst
verses, versus
vesical, vesicle

vial, vile, viol 2

vice, vise 3

vilest, violist

villain, villein (villan) 2

villainous, villainess

villous, villus

viscous, viscus

visored, vizard

vitelin, viteline

vole, wohl 2

waac, wac, wack, whack *

waacs, wacs, wacks, wax, whacks

wadder, water, watter 2

wade, weighed

waded, waited, weighted

wadi, waddy

wading, waiting, weighting 2

wail, wale, whale 3

wailer, waler, whaler 2

wain, wane 2

waist, waste 2

waister, waster 2

waistless, wasteless

wait, weight 3

waive, wave 3

waiver, waver 2

want, wont 3

war, wore

ware, wear, were, where 3

warn, worn

wart, wort 2

wary, wherry

watt, what, wot

way, weigh, whey 2

we, wee, whee

weak, week 2, 5

weal, we'll, wheal, wheel *

weald, wheeled, wield*

wealds, weals, wheals, wheels, wields

wean, ween, wheen 2

weasand, wizened

weather, whether

weave, we've

we'd, weed

weir, we're

weld, welled

welds, wells

we'll, will

welder, weldor 2

wen, when *

wends, wens, whens

wer, were, we're, where, whir(r) 2

wert, wort 2
whees, wheeze
which, wich (wych), witch 2
while, wile 3
whiled, wild, wiled
whiles, wilds, wiles
whin, win 3
whine, wine 3
whined, wind, wined
whines, winds, wines
whins, winds, wins
whirl, whorl 3
whirled, whorled, world
whirls, whorls, worlds
whirred, word
whish, wish 3
whished, whisht, wished
whist, wist
whit, wit 2
white, wight, wite 2
whither, wither
whittle, wittol 2
who, whoo 2
whoa, woe, wough 2
whoof, woof 3

why, wye, Y *

whys, wise, wyes, Y's

windlass, windless

with, withe

wood, wooed

wood, would

wreathes, wreaths

xanthene, xanthine

xanthin, xanthine

xenia, zinnia 2

xiphioid, ziphioid

ya, yah 2

yabbi, yabby 2

yahoo, yahu

y'all, yauld, yawl

yap, yapp 2

yar, yarr 2

yawed, yod(h)

yawn, yon

yawned, yond

ye, yi

yeld, yelled

yewen, yuan

yogh, yoke, yolk 2

yore, your, you're

you'll, yule

yuca, yucca 2

zacks, zax

zinc, zink 2

172 A

APPENDIX B

Pilot Study Instruments

Underline any words in the verse selection below which have homophones. (A homophone is pronounced the same as another English word, but spelled differently.)

Here then I rest. Some universal Cause
Acts to one end, though seen through various laws.
Let this great lesson guide you night and day,
But most be with you if you preach or pray.
Sword at thy side and bound in links of steel
Invade her throne, vain Reason to reveal.
Have not her ways of rigor made her wrong,
Still for the fair too foul, the weak too strong?
And thus sweet fancy gilds with morning rays
The very pride of mind that wastes our days.

Who taught whole nations of the field and wood
To flee their poison and to choose their food?
Know Nature's sons are all allowed her care;
The furs that please a Monarch warm no bear.
And lo, new needs, worn hopes, old pains must rise
Which graft belligerence on charities.
So lives that perish other forms supply:
By turns we slay the breath of life and die,
Like bubbles on the sea of matter borne,
We sail, and break, and go whence none return.

APPENDIX C

Experimental Instruments

Underline every word on the list below which is spelled differently but pronounced the same as some other English dictionary word.

lone
stole
road
flower
size
fleet

might
feet
navel
wake
pause
pour

lane
find
still
peer
tide
borders

breed
sight
colonel
lax
unreel
spoke

aloud
present
ball
loaves
quartz
mist

capital
boulder
code
roll
freight
holy

freeze
cygnet
arm
neckless
peals
prays

sell
serial
hire
cheek
oversees
chants

STOP HERE

WAIT FOR FURTHER INSTRUCTIONS

Underline any words in the verse selection below which have homophones. (A homophone is pronounced the same as another English word, but spelled differently.)

Go forth, I pray: some fair handmaid of right

Waits at one throne, all seen beneath blue night.

Lo, none need die through ways too coarse or foul;

Be made to flee him, find not horse but soul!

The lynx sees great sons perish for their wear

(Furs worn by Kings would please no grizzly bear),

So do we veil, and gild with morning rays

Pains born of pride, which waste in vain our days;

Whole motes on the rough sea of matter grown,

You sail, and break, and rest in peace unknown.

STOP HERE

WAIT FOR FURTHER INSTRUCTIONS

Underline all the words on this page which have more than one meaning.

deer
wound
fair
might
ale
desert

lead
obstacle
flatter
tablet
down
plain

thy
quarry
discuss
tart
recounted
firm

muse
brood
dual
organ
furrier
lives

content
do
batter
divers
curry
recent

jade
pail
forged
close
boring
quickly

leaves
mane
resumes
lighter
tense
rebel

dove
nail
opal
produce
read
mole

STOP HERE

WAIT FOR FURTHER INSTRUCTIONS

GPO 822-674—